

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 01:32:26 / Search time 32.1698 Seconds
(without alignments)
43.915 Million cell updates/sec

Title: US-10-088-639A-2_COPY_158_162

Perfect score: 25

Sequence: 1 SSVNH 5

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database :

A_Geneseq_29Jan04:*

1: Geneseq1980s:*

2: Geneseq1990s:*

3: Geneseq2000s:*

4: Geneseq2001s:*

5: Geneseq2002s:*

6: Geneseq2003as:*

7: Geneseq2003bs:*

8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	25	100.0	69	5	ABP33613 Human ORF
2	25	100.0	76	4	AAM57202 Human bra
3	25	100.0	76	4	ABG51282 Human liv
4	25	100.0	96	4	ABP11748 Human liv
5	25	100.0	112	4	AAE11748 Human odo
6	25	100.0	117	4	AAE11748 Human odo
7	25	100.0	117	4	AAE11748 Human odo
8	25	100.0	133	4	ABG10765 Novel hum
9	25	100.0	144	7	ADC97464 E. faeciu
10	25	100.0	167	4	ABP33613 Human ORF
11	25	100.0	248	6	ABR43245 Human PMM
12	25	100.0	249	4	ABG68087 An anti-a
13	25	100.0	253	5	ABG66935 Novel G-p
14	25	100.0	259	7	ADB47696 A. gossyp
15	25	100.0	282	7	AAO23391 Human 656
16	25	100.0	286	7	ADB47698 A. gossyp
17	25	100.0	300	5	ABP33613 Human ORF
18	25	100.0	300	6	AAE35322 Human sec
19	25	100.0	306	4	AAE11905 Human G-p
20	25	100.0	313	4	AAE11905 Human G-p
21	25	100.0	314	4	AAE11906 Human G-p
22	25	100.0	314	4	AAE11906 Human G-p
23	25	100.0	314	4	AAE11906 Human G-p
24	25	100.0	314	4	AAE11906 Human G-p
25	25	100.0	314	4	AAU24560 Human oif

26	25	100.0	314	4	AAU24558 Human oif
27	25	100.0	314	5	AAO15487 Human G-p
28	25	100.0	314	5	ABP95676 Human GPC
29	25	100.0	314	5	ABP95675 Human GPC
30	25	100.0	314	5	AAU95702 Human oif
31	25	100.0	314	5	AAU95703 Human oif
32	25	100.0	314	5	AAU85178 Human oif
33	25	100.0	314	5	AAU85180 G-coupled
34	25	100.0	314	7	ADC86303 Human GPC
35	25	100.0	314	7	ADC86305 Human GPC
36	25	100.0	316	5	ABG69806 Human REM
37	25	100.0	316	5	AAU10286 Human G p
38	25	100.0	316	6	ADA50465 Human pro
39	25	100.0	318	6	ABR01667 Human G p
40	25	100.0	321	4	AAE11903 Human G-p
41	25	100.0	322	4	AAE11904 Human G-p
42	25	100.0	352	6	ABU47907 Protein e
43	25	100.0	352	6	ABU45103 Protein e
44	25	100.0	352	6	ABU47160 Protein e
45	25	100.0	352	6	ABU28161 Protein e
46	25	100.0	353	4	ABB57766 Drosophil
47	25	100.0	353	6	ABU32217 Protein e
48	25	100.0	432	6	ABR43244 Human PMM
49	25	100.0	436	5	ABP98196 Protease-
50	25	100.0	472	4	ABB63263 Drosophil
51	25	100.0	727	2	AAO55333 Fragment
52	25	100.0	1558	4	ABG28804 Novel hum
53	24	96.0	23	2	AAE89193 GPR adren
54	24	96.0	23	2	AAE89193 GPR adren
55	24	96.0	23	2	AAE89193 GPR adren
56	24	96.0	23	2	AAE89193 GPR adren
57	24	96.0	58	4	AAE17559 Peptide #
58	24	96.0	58	4	AAE17559 Peptide #
59	24	96.0	58	4	AAE17559 Peptide #
60	24	96.0	58	4	AAE17559 Peptide #
61	24	96.0	58	4	AAE17559 Peptide #
62	24	96.0	58	4	AAE17559 Peptide #
63	24	96.0	58	4	AAE17559 Peptide #
64	24	96.0	58	4	AAE17559 Peptide #
65	24	96.0	58	4	AAE17559 Peptide #
66	24	96.0	58	4	AAE17559 Peptide #
67	24	96.0	58	4	AAE17559 Peptide #
68	24	96.0	58	4	AAE17559 Peptide #
69	24	96.0	58	4	AAE17559 Peptide #
70	24	96.0	58	4	AAE17559 Peptide #
71	24	96.0	58	4	AAE17559 Peptide #
72	24	96.0	58	4	AAE17559 Peptide #
73	24	96.0	58	4	AAE17559 Peptide #
74	24	96.0	58	4	AAE17559 Peptide #
75	24	96.0	58	4	AAE17559 Peptide #
76	24	96.0	58	4	AAE17559 Peptide #
77	24	96.0	58	4	AAE17559 Peptide #
78	24	96.0	58	4	AAE17559 Peptide #
79	24	96.0	58	4	AAE17559 Peptide #
80	24	96.0	58	4	AAE17559 Peptide #
81	24	96.0	58	4	AAE17559 Peptide #
82	24	96.0	58	4	AAE17559 Peptide #
83	24	96.0	58	4	AAE17559 Peptide #
84	24	96.0	58	4	AAE17559 Peptide #
85	24	96.0	58	4	AAE17559 Peptide #
86	24	96.0	58	4	AAE17559 Peptide #
87	24	96.0	58	4	AAE17559 Peptide #
88	24	96.0	58	4	AAE17559 Peptide #
89	24	96.0	58	4	AAE17559 Peptide #
90	24	96.0	58	4	AAE17559 Peptide #
91	24	96.0	58	4	AAE17559 Peptide #
92	24	96.0	58	4	AAE17559 Peptide #
93	24	96.0	58	4	AAE17559 Peptide #
94	24	96.0	58	4	AAE17559 Peptide #
95	24	96.0	58	4	AAE17559 Peptide #
96	24	96.0	58	4	AAE17559 Peptide #
97	24	96.0	58	4	AAE17559 Peptide #
98	24	96.0	58	4	AAE17559 Peptide #

99	22	88.0	40	4	AA023378
100	22	88.0	40	5	ABG36444
101	22	88.0	41	3	AA056046
102	22	88.0	41	4	AA089982
103	22	88.0	42	6	ABR44177
104	22	88.0	50	3	AA056045
105	22	88.0	50	5	ABP07769
106	22	88.0	56	4	AAU21988
107	22	88.0	56	5	ABP04794
108	22	88.0	56	7	AAE45956
109	22	88.0	57	4	AA087201
110	22	88.0	60	5	ABP08076
111	22	88.0	63	4	AA023516
112	22	88.0	63	4	AA023551
113	22	88.0	64	4	AAU49085
114	22	88.0	64	6	AA045604
115	22	88.0	67	5	ABP32077
116	22	88.0	71	2	AAV28807
117	22	88.0	73	4	AA074207
118	22	88.0	73	5	ABP06630
119	22	88.0	76	4	ABG27455
120	22	88.0	77	4	ABG04481
121	22	88.0	80	4	AAU32298
122	22	88.0	83	4	AA085181
123	22	88.0	83	4	ABG12693
124	22	88.0	84	4	AA091971
125	22	88.0	84	4	AA039570
126	22	88.0	84	4	AA063064
127	22	88.0	84	5	ABG97622
128	22	88.0	84	6	ABM59583
129	22	88.0	84	7	ABD92878
130	22	88.0	85	4	AA066845
131	22	88.0	86	3	AA019836
132	22	88.0	86	5	ABP34978
133	22	88.0	86	7	ADC97427
134	22	88.0	90	4	AA095433
135	22	88.0	90	4	ABE96123
136	22	88.0	96	5	ABP28326
137	22	88.0	98	2	AAU35960
138	22	88.0	98	4	AAU14406
139	22	88.0	102	3	AA053243
140	22	88.0	104	4	AA090569
141	22	88.0	104	5	ABP42981
142	22	88.0	108	4	AAU17180
143	22	88.0	108	4	ABG01082
144	22	88.0	108	7	ABD93888
145	22	88.0	110	4	ABE11061
146	22	88.0	110	4	ABE17935
147	22	88.0	116	4	AAU31394
148	22	88.0	128	4	AAO07167
149	22	88.0	128	4	AAO01243
150	22	88.0	135	4	AAO01920

ALIGNMENTS

RESULT 1
 ID ABP33613 standard; protein; 69 AA.
 AC ABP33613;
 DT 08-JUL-2002 (first entry)
 DE Human ORF2586 protein, SEQ ID NO:5172.
 KW Human; ORF; open reading frame; ORFX; drug screening; diagnosis;
 KW disease monitoring; cytokine; cell proliferation; cell differentiation;
 KW immune modulation; haematopoiesis regulation; tissue growth;
 KW angiogenesis; activin; inhibitor; chemotactic; chemokinetic; haemostatic;
 KW thrombolytic; tumour inhibition; bodily characteristic; fertility;
 KW behaviour; cancer; proliferative disorder; neurological disorder;

KW cardiovascular disease; immune system disorder; organ transplantation;
 KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
 KW hypothyroidism; cholesterol ester storage disease; infection; vulnery;
 KW vasotropic; antipsoriatic; antidiabetic; cytostatic; neutropenic;
 KW neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;
 KW cardiant; hypotensive; antithyroid; antiinflammatory; immunomodulator;
 KW dermatological; analgesic; virucide; antibacterial; fungicide.
 XX Homo sapiens.
 OS WO200190366-A2.
 XX 29-NOV-2001.
 XX 24-MAY-2001; 2001WO-US017076.
 XX 24-MAY-2000; 2000US-0206690P.
 XX (CURA-) CURAGEN CORP.
 XX Leach MD, Shinkets RA;
 XX WPI; 2002-106200/14.
 XX N-PSDB; ABN77639.
 XX Novel human polypeptides and polynucleotides useful for diagnosing,
 XX preventing and treating cardiovascular disease, neurodegenerative,
 XX hyperproliferative disorders and disorders related to organ
 XX transplantation.
 XX Claim 10; Page 1545; 2508pp; English.
 XX Sequences ABP31028-ABP35561 represent 4534 novel human proteins
 XX designated ORF (open reading frame) 1-4534, and sequences ABN75054-
 XX ABN79587 represent cDNAs encoding them. The invention also encompasses
 XX polypeptides at least 80% identical to the ORF1-ORF4534 (collectively
 XX referred to as ORFX) proteins, polynucleotides at least 85% identical to
 XX the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
 XX polynucleotides, the recombinant production of ORFX proteins, antibodies
 XX specific for ORFX proteins, methods of detecting ORFX expression or
 XX polypeptides, and methods of screening individuals for a predisposition to an
 XX ORFX-associated disorder. The ORFX proteins of the invention have a wide
 XX range of biological activities, such as cytokine, cell proliferation,
 XX cell differentiation, immune modulation, haematopoiesis regulation,
 XX tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
 XX chemokinetic activity, haemostatic activity, thrombolytic activity,
 XX receptor/ligand, antiinflammatory activity, tumour inhibition activity,
 XX of bodily characteristic activity, and may also be involved in the determination
 XX of other proliferative disorders such as psoriasis and benign tumours,
 XX neurological disorders such as epilepsy and Alzheimer's disease,
 XX cardiovascular diseases, immune system disorders, disorders related to
 XX organ transplantation, disorders of tissue growth and regeneration,
 XX diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester
 XX storage disease, and infectious diseases caused by viral, bacterial,
 XX fungal and other pathogens. ORFX nucleic acids may also be used as a
 XX source of primers and probes, in the detection of ORFX genomic sequences
 XX or transcripts, in the identification and cloning of homologous
 XX sequences, in genetic diagnosis, and in forensic biology. The ORFX
 XX nucleic acids may additionally be used to produce transgenic animals
 XX which may be useful for studying the function and/or activity of ORFX
 XX protein, and in drug screening. The ORFX proteins may also be used as
 XX immunogens to generate specific antibodies, which are useful in the
 XX diagnosis, treatment and monitoring of ORFX-associated diseases
 XX

Sequence 69 AA;
 Query Match 100.0%; Score 25; DB 5; Length 69;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 SSVMH 5
Db 23 SSVMH 27

RESULT 2
AAM57202
ID AAM57202 standard; protein; 76 AA.
XX
AC AAM57202;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 29307.
XX
KW Human; brain expressed exon; gene expression analysis; probe; microarray;
KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.
XX
OS Homo sapiens.
XX
PN WO200157275-A2.
PD
XX 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000667.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483446/52.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT brains.
XX
PS Example 4; SEQ ID NO 29307; 650pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is a protein encoded by one of
CC the probes of the invention
XX
SQ Sequence 76 AA;

Query Match 100.0%; Score 25; DB 4; Length 76;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
Db 35 SSVMH 39

RESULT 3
ABG51282
ID ABG51282 standard; peptide; 76 AA.
XX
AC ABG51282;
XX
DT 25-FEB-2003 (first entry)
XX
DE Human liver peptide, SEQ ID NO 29930.
```

```
XX
KW Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
KW hypercholesterolaemia; coronary heart disease.
XX
OS Homo sapiens.
XX
PN WO200157273-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000664.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488898/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human adult liver.
XX
PS Claim 27; SEQ ID NO 29930; 658pp; English.
XX
CC The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC measuring human gene expression in a sample derived from human adult
CC liver, comprising one of 13109 defined nucleotide sequences given in the
CC specification (or complements/ fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult liver.
CC (I) may be used for predicting, measuring and displaying gene expression
CC in samples derived from human adult liver. The genes identified may be
CC involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC associated with coronary heart disease. ABG47348-ABG59930 represent human
CC liver single exon encoded peptides of the invention. Note: The sequence
CC information for this patent does not appear in the printed specification
CC but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 76 AA;

Query Match 100.0%; Score 25; DB 4; Length 76;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
Db 35 SSVMH 39

RESULT 4
ABB11748
ID ABB11748 standard; peptide; 96 AA.
XX
AC ABB11748;
XX
DT 11-JAN-2002 (first entry)
XX
DE Human odorant receptor S18 homologue, SEQ ID NO:2118.
XX
KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW chronic inflammatory condition; proliferative retinopathy;
```

KW atherosclerosis; coronary heart disease; arterial ischaemia;
 KW bone disorder; osteoporosis; vascular growth disorder;
 KW tissue regeneration; wound healing; infection; immune disorder;
 KW cell culture; drug screening; gene therapy; antiinflammatory;
 KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
 KW cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
 KW antifungal; vulnery; antiulcer.
 XX
 OS Homo sapiens.
 XX
 XX WO200157188-A2.
 XX
 XX PD 09-AUG-2001.
 XX
 XX PF 05-FEB-2001; 2001WO-US003800.
 XX
 XX PR 03-FEB-2000; 2000US-00496914.
 XX PR 27-APR-2000; 2000US-00560875.
 XX
 XX PA (HYSE-) HYSEQ INC.
 XX
 XX PI Tang YT, Liu C, Drmanac RT;
 XX
 XX WPI; 2001-457740/49.
 DR N-PSDB; ABA08992.
 DR
 XX
 PT Human proteins and DNA encoding sequences useful for preventing, treating
 PT or ameliorating a medical condition in a mammalian subject e.g. arthritis
 PT and cancer.
 PT
 XX
 XX Claim 20; Page 238; 1963pp; English.
 XX
 XX Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
 CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
 CC invention also relates to vectors and recombinant host cells comprising a
 CC nucleotide of the invention, methods of producing the novel polypeptides,
 CC antibodies against the polypeptides, methods of detecting the nucleotides
 CC or polypeptides in a sample, and methods of identifying compounds which
 CC bind to polypeptides of the invention. Although novel, many of the
 CC polypeptides of the invention have homology to known proteins, thereby
 CC giving an insight into their probable biological activities, and hence
 CC potential therapeutic applications. The polypeptides of the invention may
 CC have various activities, including cytokine, cell proliferation or cell
 CC differentiation activities; stem cell growth factor activity;
 CC haematopoiesis regulatory activity; tissue growth activity;
 CC immunomodulatory activity; activin- or inhibin-related activities;
 CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
 CC thrombolytic activities; receptor or ligand activities; or may be
 CC involved in oncogenesis, cancer cell proliferation or metastasis.
 CC Depending on their biological activities, polypeptides and nucleotides of
 CC the invention are useful for preventing, treating or ameliorating medical
 CC conditions, e.g., by protein or gene therapy. Such conditions include
 CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
 CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
 CC proliferative retinopathy, atherosclerosis, coronary heart disease,
 CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
 CC vascular growth. Polypeptides involved with tissue regeneration and
 CC repair (or nucleic acids encoding them) may be used to promote wound
 CC healing (e.g., of burns, incisions and ulcers), while those with
 CC immunomodulatory activities may be used in the treatment of viral,
 CC bacterial and fungal infections in addition to immune disorders.
 CC Polypeptides with growth factor activity may be used in cell cultures to
 CC promote cell growth. For example, such polypeptides may be used to
 CC manipulate stem cells in culture to give rise to neuroepithelial cells
 CC that can be used to augment or replace cells damaged by illness,
 CC autoimmune disease or accidental damage. The polypeptides and nucleotides
 CC may also be used in the diagnosis of the above conditions, and in drug
 CC screening techniques. The present sequence represents a novel human
 CC polypeptide of the invention
 XX
 SQ Sequence 96 AA;

Query Match 100.0%; Score 25; DB 4; Length 96;

Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SSVMH 5
 DB 39 SSVMH 43
 |||||
 RESULT 5
 AAE12026
 ID AAE12026 standard; protein; 112 AA.
 XX
 XX AC AAE12026;
 XX
 XX DT 18-DEC-2001 (first entry)
 XX
 XX DE Human G-protein coupled receptor, GCRC-5.
 XX
 KW Human; G-protein coupled receptor; GCRC; gene therapy;
 KW proteome analysis; cell proliferative disorder; arteriosclerosis;
 KW cancer; neurological disorder; Huntington's disease; Parkinson's disease;
 KW cardiovascular disorder; atherosclerosis; congestive heart failure;
 KW gastrointestinal disorder; gastritis; nausea; autoimmune; anaemia;
 KW inflammatory disorder; acquired immunodeficiency syndrome; AIDS;
 KW metabolic disorder; diabetes; obesity; viral infection; drug screening;
 KW chromosomal mapping; cytostatic; immunomodulatory; antiinflammatory;
 KW virucide.
 XX
 XX OS Homo sapiens.
 XX
 XX FH Key Location/Qualifiers
 FT Peptide 1..27
 FT /label= Signal_peptide
 FT Domain 1..23
 FT /label= Transmembrane_domain
 FT Protein 28..112
 FT /note= "Mature human GCRC-5 protein"
 FT Binding-site 64..112
 FT /note= "Visual pigments retinal binding site"
 FT Domain 81..97
 FT /label= G-protein_coupled_receptor_domain
 XX WO200172836-A2.
 XX
 XX PD 04-OCT-2001.
 XX
 XX PF 29-MAR-2001; 2001WO-US010436.
 XX
 XX PR 29-MAR-2000; 2000US-0193051P.
 XX PR 06-APR-2000; 2000US-0195155P.
 XX PR 20-APR-2000; 2000US-0199084P.
 XX PR 28-APR-2000; 2000US-0200551P.
 XX PR 05-MAY-2000; 2000US-0202278P.
 XX
 XX (INCY-) INCYTE GENOMICS INC.
 XX
 XX PI Policky JL, Tribouley CM, Tang YT, Baughn MR, Graul R, Khan FA;
 XX Nguyen DB, Patterson C, Lal P, Au-Young J, Yang J, Hafalia A;
 XX Walia NK, Das D;
 XX
 XX WPI; 2001-616472/71.
 DR N-PSDB; AAD19581.
 DR
 XX New human G-protein coupled receptors, useful for treatment and diagnosis
 PT of e.g. cell proliferation, also screening for specific modulators, and
 PT related nucleic acid.
 XX
 XX Claim 1; Page 106; 111pp; English.
 PS
 XX The present sequence is a human G-protein coupled receptor, GCRC-5
 CC protein. The GCRCs are used for treating or preventing disorders
 CC associated with decreased expression of functional GCRC, and for
 CC identifying specific agonists and antagonists, also binding agents and

modulators. They can also be used for generating specific antibodies and for proteosome analysis. Disorders that can be treated include cell proliferative disorders, e.g., arteriosclerosis and cancer, neurological disorders, e.g., Huntington's disease and Parkinson's disease, cardiovascular disorders, e.g., atherosclerosis and congestive heart failure, gastrointestinal disorders, e.g., gastritis and nausea, autoimmune/inflammatory disorders, e.g., acquired immunodeficiency syndrome (AIDS) and anaemia, metabolic disorders, e.g., diabetes and obesity and viral infections. Nucleic acids that encode GCRC are used for identifying agents that alter its expression, for assessing toxicity of test compounds and as sources of primers and probes for diagnostic. detection of GCRC DNA and of therapeutic antisense and ribozyme sequences. They can also be used in gene therapy, for chromosomal mapping, and for recombinant production of GCRC. The antibodies are useful for diagnosis and monitoring of diseases associated with GCRC expression, for detecting and purifying GCRC, and as therapeutic agents and for drug screening

XX Sequence 112 AA;

Query Match 100.0%; Score 25; DB 4; Length 112;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSMVH 5
|||||
Db 55 SSMVH 59

RESULT 6

AAG72494
ID AAG72494 standard; protein; 117 AA.

AC AAG72494;

DT 30-JUL-2001 (first entry)

DE Human OR-like polypeptide query sequence, SEQ ID NO: 2175.

XX Human; olfactory receptor; OR; primary scent determination;

KW secondary scent determination; polypeptide library; odour receptor;

KW scent profile; scent fingerprint; scent representation.

XX Homo sapiens.

PN WO200127158-A2.

XX 19-APR-2001.

PF 06-OCT-2000; 2000WO-US027582.

XX 08-OCT-1999; 99US-0158615P.

PR 24-FEB-2000; 2000US-0184809P.

XX (DIGI-) DIGISCENTS.

PA (YEDA) YEDA RES & DEV CO LTD.

XX Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;

PI WPI; 2001-290713/30.

XX New polynucleotides which encode polypeptides involved in olfactory

PT sensation for identifying olfactory agonists and antagonists.

XX Example 6; Page 1464-1465; 1857pp; English.

PS The present sequence is a polypeptide encoded by one of 344 newly mined
XX human genes. It was used as a query sequence in a database search of
CC olfactory receptor (OR)-like sequences. The invention relates to isolated
CC polynucleotides encoding polypeptides involved in olfactory sensation.
CC The polynucleotides can be used in screening for olfactory agonists and
CC antagonists. The methods allow for the determination of primary scents
CC and the identification of the odour receptors used to detect these

CC primary scents. The methods also enable determination of secondary scents
CC and the identification of combinations of odour receptors that are
CC involved in detecting such secondary scents. This enables the
CC construction of a scent representation (also called a scent fingerprint
CC or scent profile), which may be used to re-create and edit scents.
CC Libraries of olfactory receptors are useful for determining the
CC interaction pattern of a composition with the receptors, and can be used
CC for determining differences in the olfactory faculties of different
CC individuals

XX Sequence 117 AA;

Query Match 100.0%; Score 25; DB 4; Length 117;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSMVH 5
|||||
Db 63 SSMVH 67

RESULT 7

AAG71795
ID AAG71795 standard; protein; 117 AA.

AC AAG71795;

DT 30-JUL-2001 (first entry)

DE Human olfactory receptor polypeptide, SEQ ID NO: 1476.

XX Human; olfactory receptor; OR; primary scent determination;

KW secondary scent determination; polypeptide library; odour receptor;

KW scent profile; scent fingerprint; scent representation.

XX Homo sapiens.

PN WO200127158-A2.

XX 19-APR-2001.

PF 06-OCT-2000; 2000WO-US027582.

XX 08-OCT-1999; 99US-0158615P.

PR 24-FEB-2000; 2000US-0184809P.

XX (DIGI-) DIGISCENTS.

PA (YEDA) YEDA RES & DEV CO LTD.

XX Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;

PI WPI; 2001-290713/30.

XX New polynucleotides which encode polypeptides involved in olfactory
PT sensation for identifying olfactory agonists and antagonists.

XX Claim 11; Page 934-935; 1857pp; English.

CC The present sequence is an olfactory receptor which is encoded by one of
CC a number of novel polynucleotides. The polynucleotides can be used in
CC screening for olfactory agonists and antagonists. The methods allow for
CC the determination of primary scents and the identification of the odour
CC receptors used to detect these primary scents. The methods also enable
CC determination of secondary scents and the identification of combinations
CC of odour receptors that are involved in detecting such secondary scents.
CC This enables the construction of a scent representation (also called a
CC scent fingerprint or scent profile), which may be used to re-create and
CC edit scents. Libraries of olfactory receptors are useful for determining
CC the interaction pattern of a composition with the receptors, and can be
CC used for determining differences in the olfactory faculties of different
CC individuals

XX Sequence 117 AA;

Query Match 100.0%; Score 25; DB 4; Length 117;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
 |||||
 DB 63 SSVMH 67

RESULT 8
 ABG10765
 ID ABG10765 standard; protein; 133 AA.

XX AC ABG10765;

XX DT 13-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #10756.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

DR N-PSDB; AAS74952.

XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.

PS Claim 20; SEQ ID NO 41124; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 133 AA;

Query Match 100.0%; Score 25; DB 4; Length 133;

Best Local Similarity 100.0%; Pred. No. 2.2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
 |||||
 DB 12 SSVMH 16

RESULT 9

ADC97464

ID ADC97464 standard; protein; 144 AA.

XX AC ADC97464;

XX DT 01-JAN-2004 (first entry)

XX DE E. faecium protein sequence SEQ ID 7091.

XX KW Vaccine; urinary tract infection; bacteraemia; endocarditis; wound;
 KW abdominal-pelvic infection.

XX OS Enterococcus faecium.

XX PN US6583275-B1.

XX PD 24-JUN-2003.

XX PF 30-JUN-1998; 98US-00107532.

XX PR 02-JUL-1997; 97US-0051571P.

XX PR 14-MAY-1998; 98US-0085598P.

XX PA (GENO-) GENOME THERAPEUTICS CORP.

XX PI Doucette-Stamm LA, Bush D;

XX WPI; 2003-799836/75.

DR N-PSDB; ADC93810.

XX New isolated nucleic acid derived from Enterococcus faecium encoding an
 PT Enterococcus faecium polypeptide useful for detection, prevention and
 PT treatment of a pathological condition resulting from a bacterial
 PT infection.

PS Example 1; SEQ ID NO 7091; 243pp; English.

XX The invention relates to an isolated nucleic acid derived from
 CC Enterococcus faecium encoding an Enterococcus faecium polypeptide having
 CC one of 10 fully defined sequences given in the (or comprising 40
 CC sequential nucleotides chosen from any of the nucleic acids, its
 CC complement or sequences hybridising to it). Also included are a
 CC recombinant vector comprising the nucleic acid operably linked to a
 CC transcription regulatory element, a cell comprising the vector and a
 CC single-stranded probe comprising the nucleic acid. The nucleic acids are
 CC chosen from 3654 disclosed sequences encoding 3654 disclosed proteins.
 CC The nucleic acids is useful for diagnosing pathological conditions
 CC resulting from E. faecium bacterial infection (e.g. urinary tract
 CC infection, bacteraemia, endocarditis, wounds and abdominal-pelvic
 CC infection) and for screening drugs such as agonists and antagonists. The
 CC nucleic acid is useful for recombinant production of Candida albicans -
 CC derived peptides or antisense polypeptides. Pharmaceutical compositions
 CC and vaccines containing the nucleic acid are useful for preventing or
 CC treating Enterococcus faecium infections. The present sequence represents
 CC one if the disclosed E. faecium proteins.

XX Sequence 144 AA;

Query Match

Best Local Similarity 100.0%; Score 25; DB 7; Length 144;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
 |||||

Db 94 SSMH 98

RESULT 10
AAB85005

ID AAB85005 standard; protein; 167 AA.
XX
AC AAB85005;
XX
DT 06-AUG-2001 (first entry)
XX
DE Amino acid sequence of A1138218, a PHOR-1 family member.
XX
KW G-protein-coupled receptor; prostate; cancer; PHOR-1; kidney; uterine;
KW cervical; stomach; rectal; cytostatic; vaccine; cell function regulator;
KW human; prostate homologue of olfactory receptor-1.
XX
OS Homo sapiens.
XX

Key Location/Qualifiers
FH Misc-difference 1..167
FT /note= "residues Xaa are encoded by internal stop codons"
XX
XX
PN WO200125434-A1.
XX
PD 12-APR-2001.
XX
XX
PF 05-OCT-2000; 2000WO-US027543.
XX
PR 05-OCT-1999; 99US-0157902P.
XX
PA (UROC-) UROGENESYS INC.
XX
PI Raitano AB, Afar DEH, Jakobovits A, Paris M, Hubert RS;
PI Mitchell SC, Safran DC;
XX
DR WPI; 2001-367230/38.
DR N-PSDB; AAF83882.
XX
XX
PT Novel gene designated PHOR-1, a G-protein-coupled receptor up-regulated
PT in prostate cancer, useful as diagnostic marker and therapeutic target
PT for cancers of prostate, kidney, uterus.
XX
PS Disclosure; Fig 22; 139pp; English.
XX
XX
CC The invention relates to a novel G-protein-coupled receptor up-regulated
CC in prostate cancer, termed PHOR-1. The encoding cDNA is contained in
CC plasmid designated p101PAl1 deposited with ATCC as Accession No. PTA-312.
CC PHOR-1 polypeptides and polynucleotides are useful for diagnosing the
CC presence of cancer, especially prostate, kidney, uterine, cervical,
CC stomach or rectal cancer by determining and comparing the level of the
CC protein or mRNA expression in test and normal tissue samples.
CC Pharmaceutical compositions comprising PHOR-1 is useful for treating
CC cancer. PHOR-1 proteins are useful for identifying ligands and other
CC agents and cellular constituents that binds to PHOR-1 gene product and
CC for generating antibodies which are useful in diagnostic, prognostic and
CC imaging methodologies and for the treatment of prostate cancer. Cell
CC lines expressing PHOR-1 are useful for identifying protein-protein
CC interactions mediated by PHOR-1. The present sequence represents the
CC amino acid sequence of A1138218, a PHOR-1 family member
XX
SQ Sequence 167 AA;

Query Match 100.0%; Score 25; DB 4; Length 167;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSMH 5
|
|
|
|
Db 58 SSMH 62

RESULT 11

ABR43245
ID ABR43245 standard; protein; 248 AA.
XX
AC ABR43245;
XX
DT 07-JUL-2003 (first entry)
XX
DE Human PMMM-6 protein SEQ ID NO:6.
XX
KW Human; protein modification and maintenance molecule; PMMM; cytostatic;
KW antiarteriosclerotic; anticonvulsant; nootropic; neuroprotective; AIDS;
KW cerebroprotective; anti-HIV; antiallergic; antiinflammatory; cancer;
KW thrombolytic; gene therapy; cell proliferative disorder; atherosclerosis;
KW neurological disorder; epilepsy; Huntington's disease; stroke; allergy;
KW immune disorder; inflammatory disorder; developmental disorder;
KW hypothyroidism; Cushing's syndrome; infection.
XX
OS Homo sapiens.
XX
PN WO2003025131-A2.
XX
PD 27-MAR-2003.
XX
PF 13-SEP-2002; 2002WO-US029221.
XX
XX
PR 14-SEP-2001; 2001US-0322196P.
PR 21-SEP-2001; 2001US-0324134P.
PR 05-OCT-2001; 2001US-0327233P.
PR 26-OCT-2001; 2001US-0346198P.
PR 02-NOV-2001; 2001US-0343980P.
PR 09-NOV-2001; 2001US-0348887P.
PR 16-NOV-2001; 2001US-0332423P.
PR 28-NOV-2001; 2001US-0334145P.
PR 28-NOV-2001; 2001US-0334229P.
PR 06-DEC-2001; 2001US-0337451P.
PR 25-JAN-2002; 2002US-0351928P.
PR 21-MAR-2002; 2002US-0366837P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
XX
PI Sprague WM, Chawla NK, Warren BA, Tang YT, Elliott VS;
PI Marquis JP, Li JX, Griffin JA, Gietzen KJ, Yang J, Lu DAM;
PI Emerling BM, Duggan BM, Richardson TW, Lee SY, Ramkumar J, Becha SD;
PI Lehr-Mason PM, Swarnakar A, Tran UK, Kable AE, Hafalia AJA, Khare R;
XX
XX
DR WPI; 2003-354597/33.
DR N-PSDB; ACC59964.
XX
XX
PT New human protein modification and maintenance molecules (PMMM), useful
PT for diagnosing, treating and preventing diseases or conditions associated
PT with the aberrant PMMM expression e.g. cancer, AIDS, epilepsy, or
PT infections.
XX
XX
PS Claim 1; Page 213-214; 270pp; English.
XX
XX
CC ACC59959 to ACC59989 encode the human protein modification and
CC maintenance molecule proteins given in ABR43240 to ABR43270, designated
CC PMMM-1 to PMMM-31 (1). (1) have cytostatic, antiarteriosclerotic,
CC anticonvulsant, nootropic, neuroprotective, cerebroprotective, anti-HIV,
CC antiallergic, antiinflammatory and thymimetic activities, and can be
CC used in gene therapy. The PMMM polypeptides and polynucleotides are
CC useful in diagnosing, treating and preventing diseases or conditions
CC associated with the decreased expression or overexpression of PMMM, such
CC as cell proliferative (e.g. cancer, atherosclerosis), neurological (e.g.
CC epilepsy, Huntington's disease, stroke), immune/inflammatory (e.g. AIDS,
CC allergies) and developmental (e.g. hypothyroidism, Cushing's syndrome)
CC disorders, or infections. They are also useful in assessing the effects
CC of exogenous compounds on the expression of nucleic acid and amino acid
CC sequences of PMMM. The PMMMs or their fragments are useful in screening
CC compounds for effectiveness as agonist or antagonist of the polypeptides,
CC or in altering the expression of the target polynucleotide and compounds
CC that specifically bind to or modulate the activity of the polypeptide.
CC The microarray is useful in monitoring or measuring protein-protein

CC interactions, drug-target interactions, and gene expression profiles

XX Sequence 248 AA;

Query Match 100.0%; Score 25; DB 6; Length 248;
 Best Local Similarity 100.0%; Pred. No. 4.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
 |||||

Db 234 SSVMH 238

RESULT 12

AA68087
 ID AAB68087 standard; protein; 249 AA.

AC AAB68087;

DT 09-JUL-2001 (first entry)

XX An anti-alpha6beta4 integrin light chain linked to a heavy chain.

DE Gastrointestinal epithelial tumour cell; alpha6beta4 integrin;

XX tumour-associated antigen; metastatic disease; malignant disease;

KW tumour typing; tumour screening; tumour.

XX Synthetic.

OS Macaca fascicularis.

OS Location/Qualifiers

FT Key

FT Region

FT /note="Complementarity determining region (CDR) 1 of the light chain"

FT 49..55

FT /note="Complementarity determining region (CDR) 2 of the light chain"

FT 88..98

FT /note="Complementarity determining region (CDR) 3 of the light chain"

FT 110..127

FT /note="linker"

FT 158..162

FT /note="Complementarity determining region (CDR) 1 of the heavy chain"

FT 177..193

FT /note="Complementarity determining region (CDR) 1 of the heavy chain"

FT 226..238

FT /note="Complementarity determining region (CDR) 1 of the heavy chain"

XX WO200130854-A2.

PN 03-MAY-2001.

XX 26-OCT-2000; 2000WO-SE002082.

XX 28-OCT-1999; 99SE-00003895.

XX (ACTI-) ACTIVE BIOTECH AB.

XX Brodin TN, Karlstroem PU, Ohlsson LG, Tordsson MJ, Kearney PP;

PI Nilsson BHK;

XX WPI; 2001-308619/32.

DR N-PSDB; AAP84797.

XX Novel antibody for diagnosis, treatment of human metastatic and malignant diseases, has binding structure for target structure displayed on cell surface of human gastrointestinal epithelial tumor and normal cells.

XX Claim 1; Page 55-56; 75pp; English.

XX The present sequence represents a Monkey antibody light chain linked to a heavy chain. The antibody binds to a target structure displayed in and on the cell surface of human gastrointestinal epithelial tumour cells and in a subpopulation of normal human gastrointestinal epithelial cells. The target structure especially comprises alpha6beta4 integrin. This is a tumour-associated antigen. The antibody and its fragments, are useful for treating conditions based on anti-angiogenic mechanism in humans, for example, for treating human metastatic and malignant disease, for example, in vivo diagnosis and prognosis of human malignant disease, comprising tumour typing, tumour screening, tumour diagnosis and prognosis and monitoring premalignant conditions. Quantitative in vivo diagnosis is carried out by determining the localization of antibody to tumour deposits in humans

XX Sequence 249 AA;

Query Match 100.0%; Score 25; DB 4; Length 249;
 Best Local Similarity 100.0%; Pred. No. 4.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
 |||||

Db 158 SSVMH 162

RESULT 13

ABG66935
 ID ABG66935 standard; protein; 253 AA.

AC ABG66935;

XX 24-SEP-2002 (first entry)

DT Novel G-protein coupled receptor related protein #12.

DE G protein coupled receptor; GPCR; olfactory receptor;

XX cell signal processing disorder; metabolic pathway modulation;

KW cardiomyopathy; atherosclerosis; diabetes; developmental disease;

KW immune disease; taste disorder; scent detectability disorder; obesity;

KW Burkitt's lymphoma; corticosteroid disease; infectious disease; pain;

KW signal transduction pathway disorder; metabolic pathway disorder;

KW retinal disease; metabolic disorder; cancer; Parkinson's disease;

KW acute heart failure; urinary retention; osteoporosis; Crohn's disease;

KW ulcer; allergy; neurological disorder; genetic disorder; transplantation;

KW fertility; Pancreatitis; Hyperthyroidism; Endometriosis;

XX forensic biology; transgenic animal.

XX Mus musculus.

XX WO200240539-A2.

PN 23-MAY-2002.

XX 16-OCT-2001; 2001WO-US032256.

XX 16-OCT-2000; 2000US-0240704P.

PR 26-OCT-2000; 2000US-0243497P.

PR 31-OCT-2000; 2000US-0244542P.

PR 03-NOV-2000; 2000US-0245484P.

PR 12-DEC-2000; 2000US-0255017P.

PR 17-JAN-2001; 2001US-0262159P.

PR 22-JAN-2001; 2001US-0263218P.

PR 25-JAN-2001; 2001US-0264118P.

PR 12-FEB-2001; 2001US-0268225P.

PR 15-FEB-2001; 2001US-0269031P.

PR 27-JUL-2001; 2001US-0308203P.

XX (CURA-) CURAGEN CORP.

PA Kekuda R, Spytek KA, Casman SJ, Zerhusen BD, Li L, Tchernev VT;

PI Colman SD, Ballinger RA, Padigaru M, Wolenc AR, Shenoy SG;

PI Edinger SR, Gerlach V, Gangolli EA, Macdougall JR, Smithson G;
 PI Peyman JA, Stone DJ, Gunther E, Ellerman K, Grosse WM, Alsobrook JP;
 XX Lpley DM, Burgess CB;
 DR WPI: 2002-500205/53.
 DR N-PSDB; ABR95453.
 XX
 XX Novel G protein coupled receptor especially olfactory receptor
 PT polypeptides and nucleic acids for diagnosing and treating
 PT atherosclerosis, cardiomyopathy and diabetes.
 XX
 XX Claim 1; Page 48; 309pp; English.
 XX
 CC The invention describes an isolated G protein coupled receptor X (GPCRI-
 CC 12) polypeptide, especially an olfactory receptor. GPCR polypeptides are
 CC useful for identifying an agent that binds to the polypeptide and for
 CC identifying a candidate substance or ligand molecules interacting with an
 CC olfactory receptor polypeptide. The polypeptide, (I) and (II) are also
 CC useful for treating diseases and disorders related to cell signal
 CC processing and metabolic pathway modulation e.g. cardiomyopathy,
 CC atherosclerosis and diabetes, and developmental diseases, immune
 CC diseases, taste and scent detectability disorders, Burkitt's lymphoma,
 CC corticosterogenic disease, signal transduction pathway disorders,
 CC metabolic pathway disorders, retinal diseases, metabolic disorders,
 CC obesity, infectious disease, pain, cancer, Parkinson's disease, acute
 CC heart failure, urinary retention, osteoporosis, Crohn's disease, ulcers,
 CC allergies, neurological disorders, genetic disorders, transplantation,
 CC fertility, pancreatitis, Hyperthyroidism and Endometriosis. GPCR
 CC sequences are also useful for identifying a cell or tissue type in a
 CC biological sample, to amplify DNA sequences from very small biological
 CC samples such as tissues e.g. hair or skin or body fluids in forensic
 CC biology. Cells comprising (I) are useful for producing non-human
 CC transgenic animals for studying the function and/or activity of GPCR
 CC protein and for identifying and/or evaluating modulators of GPCR protein
 CC activity. This is the amino acid sequence of a novel G-protein coupled
 CC receptor described in the invention
 XX
 SQ Sequence 253 AA;
 Query Match 100.0%; Score 25; DB 5; Length 253;
 Best Local Similarity 100.0%; Pred. No. 4.4e+02; Mismatches 0; Indels 0; Gaps 0;
 Matches 5; Conservative 0;
 QY 1 SSVMH 5
 Db 199 SSVMH 203
 |||||
 RESULT 14
 ADB47696
 ID ADB47696 standard; protein; 259 AA.
 XX
 AC ADB47696;
 XX
 DT 04-DEC-2003 (first entry)
 XX
 DE A. gossypii cell wall/cytoskeleton oligo 103 protein fragment SEQ ID 18.
 XX
 KW cell wall construction; cytoskeleton construction; vitamin B2;
 KW food supplement; Oligo 103.
 XX
 OS Eremothecium gossypii.
 XX
 PN WO2003018626-A2.
 XX
 PD 06-MAR-2003.
 XX
 PF 21-AUG-2002; 2002WO-EP009355.
 XX
 PR 22-AUG-2001; 2001DE-01041057.
 PR 22-AUG-2001; 2001DE-01041059.
 PR 22-AUG-2001; 2001DE-01041060.
 PR 22-AUG-2001; 2001DE-01041061.

PR 22-AUG-2001; 2001DE-01041063.
 PR 22-AUG-2001; 2001DE-01041064.
 PR 22-AUG-2001; 2001DE-01041065.
 PR 22-AUG-2001; 2001DE-01041066.
 PR 06-MAR-2002; 2002DE-01009827.
 PR 11-APR-2002; 2002DE-01016028.
 PR 11-APR-2002; 2002DE-01016034.
 PR 16-MAY-2002; 2002DE-01021906.
 PR 16-MAY-2002; 2002DE-01021918.
 PR 16-MAY-2002; 2002DE-01021919.
 PR 07-JUN-2002; 2002DE-01021921.
 PR 07-JUN-2002; 2002DE-01025411.
 XX
 PA (BADI) BASF AG.
 XX
 PI Karos M, Althoefer H, Kroeger B, Revuelta Doval JL;
 XX
 DR WPI; 2003-332820/31.
 DR N-PSDB; ADB47695.
 XX
 XX New polynucleotide from Ashbya gossypii, useful for increasing microbial
 PT production of Vitamin B2, encodes protein involved in construction of
 PT cell wall and cytoskeleton.
 XX
 PS Claim 7; Page 90-91; 124pp; German.
 XX
 CC This invention describes novel polynucleotides isolated from Ashbya
 CC gossypii (also known as Eremothecium gossypii), that encode a protein
 CC associated with construction of the cell wall and/or cytoskeleton.
 CC Modulating the construction of the cell wall/cytoskeleton, especially
 CC such that the cells are stabilised, improves microbial production of
 CC Vitamin B2, a human and animal food supplement. This sequence represents
 CC a fragment of Oligo 103 which encodes an A. gossypii cell
 CC wall/cytoskeleton associated protein. Oligo 103 encodes a homologue of a
 CC Saccharomyces cerevisiae actin overexpression associated protein.
 XX
 SQ Sequence 259 AA;
 Query Match 100.0%; Score 25; DB 7; Length 259;
 Best Local Similarity 100.0%; Pred. No. 4.5e+02; Mismatches 0; Indels 0; Gaps 0;
 Matches 5; Conservative 0;
 QY 1 SSVMH 5
 Db 107 SSVMH 111
 |||||
 RESULT 15
 AAO23391
 ID AAO23391 standard; protein; 282 AA.
 XX
 AC AAO23391;
 XX
 DT 06-NOV-2003 (first entry)
 XX
 DE Human 65649 metalloprotease protein.
 XX
 KW Metalloprotease; 65649; astacin; proteolytic; gene therapy; apoptosis;
 KW metastasis; wound healing; endometrial cycling; tumour cell invasion;
 KW bone remodelling; neurological; cardiovascular; endothelial; cytoskeletal;
 KW immune; bone metabolism; neuroprotective; nootropic; immunosuppressive;
 KW enzyme; human.
 XX
 OS Homo sapiens.
 XX
 PN US2003022212-A1.
 XX
 PD 30-JAN-2003.
 XX
 PF 12-JUN-2002; 2002US-00167555.
 XX
 PR 13-JUN-2001; 2001US-0297938P.
 XX

```

PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Curtis RAJ;
XX
DR WPI; 2003-584995/55.
DR N-PSDB; AAL56808, AAL56809.
XX
PT New human metalloprotease, 65649, polypeptides and polynucleotides,
PT useful for modulating e.g. tumor cell invasion or metastasis, tissue or
PT organ integrity, wound healing, endometrial cycling, hair follicle
PT cycling, or ovulation.
XX
PS Claim 17; Page 37; 44pp; English.
XX
CC This invention relates to novel human nucleic acid molecules that encode
CC a metalloprotease family member designated the 65649 protein. The 65649
CC is a member of the astacin family of proteolytic enzymes that function to
CC activate growth factors, degrade polypeptides and process extracellular
CC proteins. Typically requiring zinc for catalysis, this enzyme can be used
CC for drug screening, chromosome mapping, tissue typing and in forensic
CC biology. Furthermore, gene therapy using the 65649 molecules may be used
CC to modulate tumour cell invasion, metastasis, tissue or organ integrity,
CC wound healing, endometrial cycling, hair follicle cycling, bone
CC remodelling, ovulation, embryonic development, and apoptosis. These may
CC further be used to control or prevent e.g. cellular proliferative and/or
CC differentiative disorders, as well as neurological, cardiovascular,
CC endothelial and immune disorders, and bone metabolism disorders.
CC Compositions of 65649 can therefore be described as cytostatic,
CC neuroprotective, neurotropic and immunosuppressive. The nucleic acids,
CC proteins, and antibodies of the invention are useful in screening assays,
CC predictive medicine, and in therapeutic and prophylactic treatment. This
CC polypeptide sequence is the human 65649 metalloprotease of the invention
XX
SQ Sequence 282 AA;

Query Match 100.0%; Score 25; DB 7; Length 282;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
DB 233 SSVMH 237
|||||

RESULT 16
ADB47698
ID ADB47698 standard; protein; 286 AA.
XX
AC ADB47698;
XX
DT 04-DEC-2003 (first entry)
XX
DE A. gossypii cell wall/cytoskeleton oligo 103 protein SEQ ID 20.
XX
KW cell wall construction; cytoskeleton construction; vitamin B2;
KW food supplement; Oligo 103.
XX
OS Eremothecium gossypii.
XX
PN WO2003018626-A2.
XX
PD 06-MAR-2003.
XX
PF 21-AUG-2002; 2002WO-EF009355.
XX
PR 22-AUG-2001; 2001DE-01041057.
PR 22-AUG-2001; 2001DE-01041058.
PR 22-AUG-2001; 2001DE-01041060.
PR 22-AUG-2001; 2001DE-01041061.
PR 22-AUG-2001; 2001DE-01041063.
PR 22-AUG-2001; 2001DE-01041064.
PR 22-AUG-2001; 2001DE-01041065.
PR 22-AUG-2001; 2001DE-01041066.
XX

PR 06-MAR-2002; 2002DE-01009827.
PR 11-APR-2002; 2002DE-01016028.
PR 11-APR-2002; 2002DE-01016034.
PR 16-MAY-2002; 2002DE-01021906.
PR 16-MAY-2002; 2002DE-01021918.
PR 16-MAY-2002; 2002DE-01021919.
PR 16-MAY-2002; 2002DE-01021921.
PR 07-JUN-2002; 2002DE-01025411.
XX
XX (BADI ) BASF AG.
PA
XX Karos M, Althoefer H, Kroeger B, Revuelta Doval JL;
XX
DR WPI; 2003-332820/31.
DR N-PSDB; ADB47697.
XX
PT New polynucleotide from Ashbya gossypii, useful for increasing microbial
PT production of Vitamin B2, encodes protein involved in construction of
PT cell wall and cytoskeleton.
XX
PS Claim 7; Page 94-95; 124pp; German.
XX
CC This invention describes novel polynucleotides isolated from Ashbya
CC gossypii (also known as Eremothecium gossypii), that encode a protein
CC associated with construction of the cell wall and/or cytoskeleton.
CC Modulating the construction of the cell wall/cytoskeleton, especially
CC such that the cells are stabilised, improves microbial production of
CC Vitamin B2, a human and animal food supplement. This sequence represents
CC an A. gossypii cell wall/cytoskeleton associated protein encoded by Oligo
CC 103 which is a homologue of a Saccharomyces cerevisiae actin
CC overexpression associated protein.
XX
SQ Sequence 286 AA;

Query Match 100.0%; Score 25; DB 7; Length 286;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
DB 126 SSVMH 130
|||||

RESULT 17
ABB07687
ID ABB07687 standard; protein; 300 AA.
XX
AC ABB07687;
XX
DT 10-JUN-2002 (first entry)
XX
DE Human protease peptide.
XX
KW Human; protease peptide; structural protein; extracellular reticulum;
KW transgenic; gene therapy; enzyme.
XX
XX Homo sapiens.
XX
PN WO200216566-A2.
XX
PD 28-FEB-2002.
XX
PF 23-AUG-2001; 2001WO-US026345.
XX
PR 23-AUG-2000; 2000US-0226903P.
XX
XX (APPL-) APPLERA CORP.
XX
PI Beasley EM, Li Z;
XX
DR WPI; 2002-257912/30.
DR N-PSDB; ABA95224, ABA95225.
XX

```

PT An isolated human protease peptide, for identifying modulators of its
 PT function to treat diseases and conditions mediated by protease activity.
 PS Claim 1; Fig 2; 69pp; English.
 CC The invention relates to an isolated human protease peptide. The peptide
 CC cleaves structural proteins in the extracellular reticulum. The encoding
 CC nucleic acids are useful as primers and probes e.g. in detection assays
 CC and in the polymerase chain reaction, genotyping, antisense therapy, and
 CC gene therapy. They are also used in recombinant vectors, expressing the
 CC protease peptide, and constructing recombinant cells and animals. Vectors
 CC comprising the nucleic acids are used to transform a host cell which can
 CC generate a transgenic animal. The protease peptide, antibodies and
 CC modulators are useful for treating a disease or condition mediated by a
 CC human protease protein. The present sequence represents the human
 CC protease peptide
 XX Sequence 300 AA;
 SQ

Query Match 100.0%; Score 25; DB 5; Length 300;
 Best Local Similarity 100.0%; Pred. No. 5.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
 Db 233 SSVMH 237
 |||||

RESULT 18
 AAE35322
 ID AAE35322 standard; protein; 300 AA.
 AC AAE35322;
 XX 17-JUN-2003 (first entry)
 DT Human secreted protein.
 DE Human; secreted protein; therapeutic; drug screening; immune response;
 KW pharmacogenomic analysis; tissue typing; gene therapy.
 XX Homo sapiens.
 OS

Key Location/Qualifiers
 FH 6..26
 FT Region /label= Helix_1
 FT Modified-site /note= "N-myristoylation site"
 FT Binding-site /note= "Prokaryotic membrane lipoprotein lipid attachment site"
 FT Modified-site /note= "N-myristoylation site"
 FT Modified-site /note= "N-myristoylation site"
 FT Modified-site /note= "Casein kinase II (CK2) phosphorylation site"
 FT Modified-site /note= "N-myristoylation site"
 FT Modified-site /note= "Protein kinase C (PKC) phosphorylation site"
 FT Modified-site /note= "Protein kinase C (PKC) phosphorylation site"
 FT Modified-site /note= "Casein kinase II (CK2) phosphorylation site"
 FT Modified-site /note= "Casein kinase II (CK2) phosphorylation site"
 FT Region /label= Helix_2
 FT Modified-site /note= "N-myristoylation site"
 FT Binding-site /note= "Neutral zinc metalloproteinase zinc binding signature region"
 FT Modified-site /note= "Neutral zinc metalloproteinase zinc binding signature region"

FT Modified-site /note= "Casein kinase II (CK2) phosphorylation site"
 FT Modified-site /note= "Casein kinase II (CK2) phosphorylation site"
 FT Modified-site /note= "Protein kinase C (PKC) phosphorylation site"
 FT Modified-site /note= "N-myristoylation site"
 FT Modified-site /note= "Asn is N-glycosylated"
 FT Modified-site /note= "Casein kinase II (CK2) phosphorylation site"
 FT Modified-site /note= "N-myristoylation site"
 FT Modified-site /note= "Amidation site"
 FT Modified-site /note= "cAMP- and cGMP-dependent protein kinase phosphorylation site"
 FT WO200295049-A2.
 XX 28-NOV-2002.
 PD 13-MAY-2002; 2002WO-US015025.
 XX 21-MAY-2001; 2001US-0291940P.
 PR 13-MAY-2002; 2002US-00143575.
 XX (PEKE) PE CORP NY.
 PA Zong JM, Yan C, Beasley EM;
 PI WPI: 2003-148472/14.
 DR N-PSDB; AAD53974, AAD53975.
 XX New isolated human secreted protein useful for diagnosing or treating a
 PT disease associated with an absence of, inappropriate or unwanted
 PT expression of the secreted protein, in drug screening assays, and as
 PT reagent in biological assays.
 XX Claim 1; Page 59-60; 64pp; English.
 PS The present invention relates to novel human secreted proteins and their
 CC corresponding polynucleotides. Sequences of the invention are useful as
 CC models for the development of human therapeutic compositions, aid in the
 CC identification of therapeutic proteins and serve as targets for the
 CC development of human therapeutic agents. The peptides may be used in drug
 CC screening assays, in assays to determine the biological activity of the
 CC protein, as markers for tissues in which the corresponding protein is
 CC preferentially expressed, to raise antibodies or to elicit another immune
 CC response and to quantitatively determine the levels of the protein in
 CC biological fluids. The secreted peptides are also useful for diagnosing a
 CC disease, predisposition to a disease or treating a disorder characterised
 CC by an absence of, inappropriate or unwanted expression of the protein.
 CC The antibodies are useful in pharmacogenomic analysis, for inhibiting
 CC protein function or for tissue typing. Sequences of the invention are
 CC also useful in gene therapy. The present sequence is human secreted
 CC protein
 XX Sequence 300 AA;
 SQ

Query Match 100.0%; Score 25; DB 6; Length 300;
 Best Local Similarity 100.0%; Pred. No. 5.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
 Db 233 SSVMH 237
 |||||

RESULT 19
 AAE11905
 ID AAE11905 standard; protein; 306 AA.

XX AC AAE11905;
 XX DT 18-DEC-2001 (first entry)
 XX DE Human G-protein coupled receptor 11 (GPCR11) protein.
 XX KW Human; G-protein coupled receptor 11; GPCR11; cardiomyopathy; vaccine;
 XX KW atherosclerosis; diabetes; cardiant; cytostatic; cancer; obesity; pain;
 XX KW diabetes mellitus; anorexia; cachexia; cardiomyopathy; atherosclerosis;
 XX KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
 XX KW immune disorder; haematopoietic disorders; retinal disorder; HIV;
 XX KW human immunodeficiency virus; adenocarcinoma; bulimia; asthma; ulcer;
 XX KW angina pectoris; hypotension; hypertension; Crohn's disease; anxiety;
 XX KW multiple sclerosis; schizophrenia; dementia; mental retardation;
 XX KW gene therapy; osteoporosis; urinary retention.
 XX OS Homo sapiens.
 XX FH Key Location/Qualifiers
 XX FT Peptide 1..43
 XX FT Protein /label= Signal_peptide
 XX FT /label= Mature_GPCR11_protein
 XX PN WO200181378-A2.
 XX PD 01-NOV-2001.
 XX PF 27-APR-2001; 2001WO-US013680.
 XX PR 27-APR-2000; 2000US-0199947P.
 XX PR 27-APR-2000; 2000US-0199960P.
 XX PR 14-AUG-2000; 2000US-0275226P.
 XX PR 18-DEC-2000; 2000US-0256399P.
 XX PR 18-DEC-2000; 2000US-0256524P.
 XX PR 22-DEC-2000; 2000US-0258159P.
 XX PR 28-DEC-2000; 2000US-0258511P.
 XX PR 28-DEC-2000; 2000US-0258828P.
 XX PR 04-JAN-2001; 2001US-0259659P.
 XX PR 13-MAR-2001; 2001US-00275226.
 XX PA (CURA-) CURAGEN CORP.
 XX PI Padigaru M, Mishra V, Spytek KA, Grosse WM, Szekeres ES;
 XX PI Alsobrook JP, Burgess CB, Casman SJ, Lepley DM, Gangolli EA;
 XX PI Macdougall JR, Smithson G;
 XX DR WPI; 2001-611739/70.
 XX DR N-PSDB; AAD19142.
 XX FT G-Protein coupled receptor polypeptides and NAs useful for
 XX FT preventing, diagnosing and treating cardiomyopathy, atherosclerosis,
 XX FT cancers and diabetes.
 XX PS Claim 1; Page 43; 242pp; English.
 XX CC The present sequence is human G-protein coupled receptor-11 (GPCR-11)
 XX CC protein. GPCR protein and DNA may be used in the prevention, diagnosis
 XX CC and treatment of diseases associated with inappropriate GPCR expression;
 XX CC obesity, diabetes mellitus, anorexia, cachexia, cardiomyopathy, pain,
 XX CC atherosclerosis, neurodegenerative disorders (Alzheimer's disease,
 XX CC Parkinson's disease, Huntington's disease); bulimia, immune disorder,
 XX CC haematopoietic disorders, disorders related to cell signal processing and
 XX CC metabolic pathway modulation, retinal disorder (photoreception),
 XX CC bacterial, fungal, protozoal and viral infections (HIV); cancer (neoplasm
 XX CC adenocarcinoma); angina pectoris, hypotension, hypertension, asthma,
 XX CC Crohn's disease, multiple sclerosis, ulcers, neurological disorders
 XX CC (dementia, mental retardation, schizophrenia, anxiety); acute heart
 XX CC failure, osteoporosis, myocardial infarction and urinary retention
 XX SQ Sequence 306 AA;

Query Match 100.0%; Score 25; DB 4; Length 306;
 Best Local Similarity 100.0%; Pred. No. 5.4e+02; Indels 0; Gaps 0;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
 DB 249 SSVMH 253
 |||||

RESULT 20
 AAG71840
 ID AAG71840 standard; protein; 313 AA.
 XX AAG71840;
 AC AAG71840;
 XX 30-JUL-2001 (first entry)
 DT Human olfactory receptor polypeptide, SEQ ID NO: 1521.
 XX Human olfactory receptor; OR; primary scent determination;
 KW secondary scent determination; polypeptide library; odour receptor;
 KW scent profile; scent fingerprint; scent representation.
 XX Homo sapiens.
 OS WO200127158-A2.
 PN 19-APR-2001.
 PD 06-OCT-2000; 2000WO-US027582.
 XX 08-OCT-1999; 99US-0158615P.
 XX 24-FEB-2000; 2000US-0184809P.
 XX (DIGI-) DIGISCENTS.
 XX (YEDA) YEDA RES & DEV CO LTD.
 XX Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;
 XX WPI; 2001-290713/30.
 DR New polynucleotides which encode polypeptides involved in olfactory
 XX sensation for identifying olfactory agonists and antagonists.
 PT Claim 11; Page 971-972; 1857pp; English.
 XX The present sequence is an olfactory receptor which is encoded by one of
 CC a number of novel polynucleotides. The polynucleotides can be used in
 CC screening for olfactory agonists and antagonists. The methods allow for
 CC the determination of primary scents and the identification of the odour
 CC receptors used to detect these primary scents. The methods also enable
 CC determination of secondary scents and the identification of combinations
 CC of odour receptors that are involved in detecting such secondary scents.
 CC This enables the construction of a scent representation (also called a
 CC scent fingerprint or scent profile), which may be used to re-create and
 CC edit scents. Libraries of olfactory receptors are useful for determining
 CC the interaction pattern of a composition with the receptors, and can be
 CC used for determining differences in the olfactory faculties of different
 CC individuals
 XX SQ Sequence 313 AA;

Query Match 100.0%; Score 25; DB 4; Length 313;
 Best Local Similarity 100.0%; Pred. No. 5.5e+02; Indels 0; Gaps 0;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
 DB 259 SSVMH 263
 |||||

RESULT 21
 AAE11906

(CURA-) CURAGEN CORP.

Padigaru M, Mishra V, Spytek KA, Grosse WM, Szekeres ES;
Alabrook JP, Burgess CE, Casman SJ, Lopley DM, Gangolli EA;
Macdougall JR, Smithson G;

WPI; 2001-611739/70.
DR N-PSDB; AAD19143.

G-Protein coupled receptor polypeptides and NAS useful for
preventing, diagnosing and treating cardiomyopathy, atherosclerosis,
cancers and diabetes.

Claim 1; Page 46; 242pp; English.

The present sequence is human G-protein coupled receptor-12 (GPCR-12)
protein. GPCR protein and DNA may be used in the prevention, diagnosis
and treatment of diseases associated with inappropriate GPCR expression,
obesity, diabetes mellitus, anorexia, cachexia, cardiomyopathy, pain,
atherosclerosis, neurodegenerative disorders (Alzheimer's disease,
Parkinson's disease, Huntington's disease); bulimia, immune disorder,
CC haematopoietic disorders, disorders related to cell signal processing and
metabolic pathway modulation, retinal disorder (photoreception), cancer
bacterial, fungal, protozoal and viral infections (HIV); Cancer (neoplasm
adenocarcinoma); angina pectoris, hypotension, hypertension, asthma,
Crohn's disease, multiple sclerosis, ulcers, neurological disorders
(dementia), mental retardation, schizophrenia, anxiety); acute heart
failure, osteoporosis, myocardial infarction and urinary retention

Sequence 314 AA;

Query Match . 100.0%; Score 25; DB 4; Length 314;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0

Qy 1 SSVNH 5
|||
Db 260 SSVNH 264

RESULT 22
AAG71837
ID ID AAG71837 standard; protein; 314 AA.
XX AC AAG71837;
XX AC
DT 30-JUL-2001 (first entry)
XX DE
DE Human olfactory receptor polypeptide, SEQ ID NO: 1518.
XX Human; olfactory receptor; OR; primary scent determination;
KW secondary scent determination; polypeptide library; odour receptor;
KW scent profile; scent fingerprint; scent representation.
XX Homo sapiens.
OS
XX WO200127158-A2.
PN
PD 19-APR-2001.
XX
XX 06-OCT-2000; 2000WO-US027582.
PF
XX 08-OCT-1999; 99US-0158615P.
PR
PR 24-FEB-2000; 2000US-0184809P.
XX
XX (DIGI-) DIGISCENTS.
PA (YEDA) YEDA RES & DEV CO LTD.
PA
XX Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;
PI
XX WPI; 2001-290713/30.
DR
PT New polynucleotides which encode polypeptides involved in olfactory

PT sensation for identifying olfactory agonists and antagonists.

XX Claim 11; Page 968-969; 1857pp; English.

XX The present sequence is an olfactory receptor which is encoded by one of a number of novel polynucleotides. The polynucleotides can be used in screening for olfactory agonists and antagonists. The methods allow for the determination of primary scents and the identification of the odour receptors used to detect these primary scents. The methods also enable determination of secondary scents and the identification of combinations of odour receptors that are involved in detecting such secondary scents. This enables the construction of a scent representation (also called a scent fingerprint or scent profile), which may be used to re-create and edit scents. Libraries of olfactory receptors are useful for determining the interaction pattern of a composition with the receptors, and can be used for determining differences in the olfactory faculties of different individuals

XX Sequence 314 AA;

Query Match 100.0%; Score 25; DB 4; Length 314;

Best Local Similarity 100.0%; Pred. No. 5.5e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5

DB 260 SSVMH 264

RESULT 24

AAE06754

ID AAE06754 standard; protein; 314 AA.

XX

AC AAE06754;

XX

DT 16-OCT-2001 (first entry)

XX

DE Human G-protein coupled receptor-4 (GCRC-4) protein.

XX

KW Human; G-protein coupled receptor-4; GCRC-4; cytostatic; hepatotropic; virucide; antiinflammatory; anticonvulsant; antiemetic; neuroprotective; nootropic; cerebroprotective; hypotensive; tranquiliser; vulnerary; ophthalmological; cell proliferative disorder; actinic keratosis; anorectic; arteriosclerosis; atherosclerosis; cirrhosis; hepatitis; psoriasis; cancer; neurological disorder; stroke; Alzheimer's disease; Huntington's disease; Parkinson's disease; cardiovascular disorder; epilepsy; hypertension; varicose vein; vasculitis; dysphagia; dyspepsia; anorexia; gastrointestinal disorder; pancreatitis; autoimmune disorder; Addison's disease; Crohn's disease; acquired immune deficiency syndrome; AIDS; uveitis; infection; trauma; metabolic disorder; diabetes; obesity; osteoporosis; transgenic animal; gene therapy.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT Domain 27..42

FT Domain /label= Transmembrane_domain

FT Domain 112..128

FT Domain /label= G-protein-coupled_receptor_motif

FT Domain 199..218

FT Domain /label= Transmembrane_domain

XX

PN WO200157085-A2.

XX

PD 09-AUG-2001.

XX

PP 01-FEB-2001; 2001WO-US003455.

XX

PR 02-FEB-2000; 2000US-0180093P.

PR 11-FEB-2000; 2000US-0182045P.

XX

PA (INCY-) INCYTE GENOMICS INC.

XX

PI Baughn MR, Au-Young J, Yue H;

XX

DR WPI; 2001-488869/53.

DR N-PSDB; AAD12947.

XX

XX Novel isolated human G-protein coupled receptor useful for diagnosing, preventing and treating cell proliferative, neurological, cardiovascular, gastrointestinal, autoimmune/inflammatory and metabolic disorders.

XX Claim 1; Page 111; 138pp; English.

XX

PT sensation for identifying olfactory agonists and antagonists.

XX Claim 11; Page 970-971; 1857pp; English.

XX The present sequence is an olfactory receptor which is encoded by one of a number of novel polynucleotides. The polynucleotides can be used in screening for olfactory agonists and antagonists. The methods allow for the determination of primary scents and the identification of the odour receptors used to detect these primary scents. The methods also enable determination of secondary scents and the identification of combinations of odour receptors that are involved in detecting such secondary scents.

XX Sequence 314 AA;

Query Match 100.0%; Score 25; DB 4; Length 314;

Best Local Similarity 100.0%; Pred. No. 5.5e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5

DB 260 SSVMH 264

RESULT 23

AAAG71839

ID AAG71839 standard; protein; 314 AA.

XX

AC AAG71839;

XX

DT 30-JUL-2001 (first entry)

XX

DE Human olfactory receptor polypeptide, SEQ ID NO: 1520.

XX

KW Human; olfactory receptor; OR; primary scent determination; secondary scent determination; polypeptide library; odour receptor; scent profile; scent fingerprint; scent representation.

XX

OS Homo sapiens.

XX

PN WO200127158-A2.

XX

PD 19-APR-2001.

XX

PP 06-OCT-2000; 2000WO-US027582.

XX

PR 08-OCT-1999; 99US-0158615P.

PR 24-FEB-2000; 2000US-0184809P.

XX

PA (DIGI-) DIGISCENTS.

PA (YEDA) YEDA RES & DEV CO LTD.

XX

PI Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;

XX

DR WPI; 2001-290713/30.

XX

PT New polynucleotides which encode polypeptides involved in olfactory sensation for identifying olfactory agonists and antagonists.

XX

PS Claim 11; Page 970-971; 1857pp; English.

XX

XX The present sequence is an olfactory receptor which is encoded by one of a number of novel polynucleotides. The polynucleotides can be used in screening for olfactory agonists and antagonists. The methods allow for the determination of primary scents and the identification of the odour receptors used to detect these primary scents. The methods also enable determination of secondary scents and the identification of combinations of odour receptors that are involved in detecting such secondary scents.

CC The present sequence is human G-protein coupled receptor-4 (GCRC-4)
 CC protein. The present invention relates to GCRC protein and nucleic acids
 CC encoding them. GCRC protein, its agonist or antagonist are useful for
 CC treating diseases or conditions associated with decreased expression or
 CC overexpression of functional GCRC in a patient, where the disorder is
 CC selected from cell proliferative disorders such as actinic keratosis,
 CC arteriosclerosis, atherosclerosis, cirrhosis, hepatitis, psoriasis, and
 CC cancer, neurological disorders such as epilepsy, stroke, Alzheimer's
 CC disease, Huntington's disease, Parkinson's disease, cardiovascular
 CC disorders such as hypertension, parkinson's disease, varicose veins, gastro-
 CC intestinal disorders such as dysphagia, dyspepsia, anorexia, nausea,
 CC pancreatitis, autoimmune/inflammatory disorders such as acquired
 CC immunodeficiency syndrome (AIDS), Addison's disease, Crohn's disease,
 CC uveitis, viral, bacterial, fungal, parasitic, protozoal, helminthic
 CC infections, trauma and metabolic disorders such as diabetes, obesity,
 CC osteoporosis. GCRC proteins and their cDNAs are used to assess the
 CC effects of exogenous compounds on the expression of GCRC sequences.
 CC GCRC cDNA is useful to create knock in humanised animals (pigs) or
 CC transgenic animals (mice or rats) to model human disease, for therapeutic
 CC or diagnostic purposes, for somatic or germline gene therapy, to generate
 CC hybridisation probes useful in mapping the naturally occurring genomic
 CC sequence, and in molecular biological techniques
 XX
 SQ Sequence 314 AA;

Query Match 100.0%; Score 25; DB 4; Length 314;
 Best Local Similarity 100.0%; Pred. No. 5.5e+02; Indels 0; Gaps 0;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
 Db 260 SSVMH 264

RESULT 25

AAU24560
 ID AAU24560 standard; protein; 314 AA.

XX AC AAU24560;

DT 18-DEC-2001 (first entry)

XX DE Human olfactory receptor AOLFRA7.

XX KW Human; olfactory receptor; G protein-coupled receptor; GPCR; odourant;
 food additive; cosmetic; fragrance; pharmaceutical additive.

XX OS Homo sapiens.

XX PN WO200168805-A2.

XX PD 20-SEP-2001.

XX PF 13-MAR-2001; 2001WO-US007771.

XX PR 13-MAR-2000; 2000US-0188914P.

XX PR 24-MAR-2000; 2000US-0192033P.

XX PR 12-APR-2000; 2000US-0198474P.

XX PR 24-APR-2000; 2000US-0199335P.

XX PR 26-MAY-2000; 2000US-0207702P.

XX PR 23-JUN-2000; 2000US-0213849P.

XX PR 16-AUG-2000; 2000US-0226534P.

XX PR 07-SEP-2000; 2000US-0230732P.

XX PR 07-FEB-2001; 2001US-0268662P.

XX PA (SENO-) SENOMYX INC.

XX PI Zozulya S;

XX DR WPI; 2001-570867/64.

XX DR N-PSDB; AAS42253.

XX PT Nucleic acids encoding human olfactory G protein-coupled receptors,

PT useful for screening for compounds involved in olfactory sensation, where
 PT the compounds can be used in the food, pharmaceutical and cosmetic
 PT industries to customize odors.

XX PS Claim 60; Page 103; 319pp; English.

XX The invention relates to nucleic acids encoding human olfactory
 CC receptors, OR, (a G protein-coupled receptor, GPCR). The OR's
 CC specifically recognise molecules, odourants, that elicit specific
 CC olfactory sensation. The human olfactory receptors and polynucleotides
 CC encoding them are useful for screening a library of chemical compounds
 CC for compounds that are involved in olfactory sensation. Modulators of
 CC their activity are useful for pharmacological and genetic modulation of
 CC olfactory signalling pathways. Therefore, they can be used in the food,
 CC pharmaceutical and cosmetic industries to customise odours and
 CC fragrances. The present sequence is a human olfactory receptor of the
 CC invention

XX SQ Sequence 314 AA;

Query Match 100.0%; Score 25; DB 4; Length 314;
 Best Local Similarity 100.0%; Pred. No. 5.5e+02; Indels 0; Gaps 0;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
 Db 260 SSVMH 264

Search completed: September 24, 2004, 01:53:16
 Job time : 42.1698 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 01:53:42 ; Search time 8.49057 Seconds
(without alignments)
30.402 Million cell updates/sec

Title: US-10-088-639A-2_COPY_158_162

Perfect score: 25

Sequence: 1 SSVNH 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfiles.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	25	100.0	144	4	US-09-107-532A-7091
2	25	100.0	300	4	US-09-934-551-2
3	25	100.0	300	4	US-10-224-414-2
4	25	100.0	406	4	US-09-489-039A-12736
5	25	100.0	517	4	US-09-134-000C-4515
6	25	100.0	726	6	5208144-37
7	24	96.0	23	1	US-08-118-270-93
8	24	96.0	23	5	PCT-US93-08528-93
9	24	96.0	268	4	US-09-934-551-4
10	24	96.0	268	4	US-10-224-414-4
11	24	96.0	630	4	US-09-134-001C-4615
12	22	88.0	19	4	US-09-441-502B-6
13	22	88.0	19	4	US-09-441-502B-6
14	22	88.0	35	5	PCT-US93-08528-232
15	22	88.0	35	5	PCT-US93-08528-232
16	22	88.0	71	4	US-09-621-976-5485
17	22	88.0	86	4	US-09-107-532A-7054
18	22	88.0	201	4	US-09-328-352-5968
19	22	88.0	220	4	US-09-328-352-5307
20	22	88.0	260	2	US-07-857-224B-27
21	22	88.0	267	4	US-09-663-600A-190
22	22	88.0	279	4	US-09-489-039A-13210
23	22	88.0	306	4	US-09-252-991A-19132
24	22	88.0	316	4	US-09-134-000C-4441
25	22	88.0	318	4	US-09-252-991A-21306
26	22	88.0	323	4	US-09-489-039A-13907
27	22	88.0	339	3	US-08-688-988-2
28	22	88.0	349	4	US-09-252-991A-25984
29	22	88.0	386	3	US-09-586-719-10
30	22	88.0	399	4	US-09-506-066E-10
31	22	88.0	404	4	US-09-644-460-29
32	22	88.0	427	4	US-09-506-066E-8
33	22	88.0	451	3	US-08-969-644-8
34	22	88.0	451	3	US-08-444-189-8
35	22	88.0	451	3	US-08-468-544-8
36	22	88.0	490	4	US-08-956-171E-5212
37	22	88.0	498	4	US-09-786-240-1
38	22	88.0	501	4	US-09-252-991A-17382
39	22	88.0	506	4	US-09-134-001C-4049
40	22	88.0	548	3	US-08-687-590-31
41	22	88.0	566	2	US-08-272-255-8
42	22	88.0	566	2	US-08-484-993B-41
43	22	88.0	566	2	US-08-484-158B-41
44	22	88.0	566	2	US-08-484-56A-41
45	22	88.0	566	2	US-08-480-150A-41
46	22	88.0	566	3	US-08-458-731-41
47	22	88.0	566	3	US-08-149-223A-41
48	22	88.0	566	5	PCT-US95-08565-8
49	22	88.0	611	4	US-09-252-991A-29866
50	22	88.0	659	4	US-09-134-001C-5537
51	22	88.0	711	4	US-09-228-986-79
52	22	88.0	781	1	US-08-373-134D-2
53	22	88.0	781	2	US-09-114-637-2
54	22	88.0	787	4	US-09-328-352-8055
55	22	88.0	1182	4	US-09-287-354-5
56	22	88.0	1207	4	US-09-287-354-5
57	22	88.0	1383	4	US-09-540-236-3580
58	22	88.0	1403	4	US-09-262-537-6
59	22	88.0	1442	1	US-08-015-986A-3
60	22	88.0	1442	1	US-08-446-363-3
61	22	88.0	1445	1	US-08-015-986A-2
62	22	88.0	1445	2	US-08-446-363-2
63	22	88.0	1498	2	US-08-404-531B-28
64	22	88.0	1498	2	US-08-404-531B-29
65	22	88.0	1498	3	US-08-476-900A-28
66	22	88.0	1498	3	US-08-476-900A-29
67	22	88.0	1498	3	US-08-488-546A-28
68	22	88.0	1498	3	US-08-488-546A-29
69	22	88.0	1581	2	US-08-404-531B-6
70	22	88.0	1581	3	US-08-476-900A-6
71	22	88.0	1581	3	US-08-488-546A-6
72	22	88.0	1581	3	US-08-726-320-4
73	22	88.0	1581	3	US-09-208-716-4
74	22	88.0	1582	2	US-08-404-531B-9
75	22	88.0	1582	3	US-08-476-900A-9
76	22	88.0	1582	3	US-08-488-546A-9
77	22	88.0	1582	3	US-08-726-320-5
78	22	88.0	1582	3	US-09-208-716-5
79	22	88.0	1695	4	US-09-866-108A-15753
80	22	88.0	3170	2	US-07-642-734C-5
81	22	88.0	3170	3	US-08-439-009A-5
82	22	88.0	15281	2	US-08-471-119A-2
83	21	84.0	9	4	US-09-864-541A-43
84	21	84.0	22	1	US-08-118-270-86
85	21	84.0	22	5	PCT-US93-08528-86
86	21	84.0	35	2	US-08-637-759B-264
87	21	84.0	35	3	US-08-871-355A-264
88	21	84.0	35	4	US-09-201-945-264
89	21	84.0	37	1	US-08-118-270-227
90	21	84.0	37	5	PCT-US93-08528-227
91	21	84.0	38	1	US-08-118-270-197
92	21	84.0	38	5	PCT-US93-08528-197
93	21	84.0	43	1	US-08-118-270-198
94	21	84.0	43	5	PCT-US93-08528-198
95	21	84.0	45	1	US-08-118-270-196
96	21	84.0	45	5	PCT-US93-08528-196
97	21	84.0	46	1	US-08-118-270-199
98	21	84.0	46	5	PCT-US93-08528-199
99	21	84.0	47	1	US-08-118-270-203
100	21	84.0	47	5	PCT-US93-08528-203

101 21 84.0 48 1 US-08-118-270-200 Sequence 200, Appl
102 21 84.0 48 5 PCR-US93-08528-200 Sequence 200, Appl
103 21 84.0 56 3 US-07-741-453A-47 Sequence 47, Appl
104 21 84.0 57 3 US-07-741-453A-48 Sequence 48, Appl
105 21 84.0 57 3 US-07-741-453A-49 Sequence 49, Appl
106 21 84.0 60 4 US-09-543-681A-4304 Sequence 4304, Ap
107 21 84.0 62 4 US-09-621-976-5584 Sequence 5584, Ap
108 21 84.0 79 3 US-08-905-223-2777 Sequence 277, Appl
109 21 84.0 82 4 US-09-489-039A-14014 Sequence 14014, A
110 21 84.0 84 4 US-09-621-976-7735 Patent No. 5514582
111 21 84.0 86 6 5514582-41 Sequence 9, Appli
112 21 84.0 107 1 US-08-399-106A-9 Sequence 10, Appl
113 21 84.0 107 1 US-08-399-106A-10 Sequence 11, Appl
114 21 84.0 107 1 US-08-399-106A-11 Sequence 11, Appl
115 21 84.0 107 1 US-08-433-105A-9 Sequence 9, Appli
116 21 84.0 107 1 US-08-433-105A-10 Sequence 10, Appl
117 21 84.0 107 1 US-08-433-105A-11 Sequence 11, Appl
118 21 84.0 107 2 US-08-434-869A-9 Sequence 9, Appli
119 21 84.0 107 2 US-08-434-869A-10 Sequence 10, Appl
120 21 84.0 107 2 US-08-434-869A-11 Sequence 11, Appl
121 21 84.0 107 2 US-08-070-116A-5 Sequence 5, Appli
122 21 84.0 107 3 US-08-444-644-22 Sequence 22, Appl
123 21 84.0 107 3 US-08-444-644-39 Sequence 39, Appl
124 21 84.0 107 3 US-08-444-644-45 Sequence 45, Appl
125 21 84.0 107 4 US-08-232-246A-22 Sequence 22, Appl
126 21 84.0 107 4 US-08-232-246A-39 Sequence 39, Appl
127 21 84.0 107 4 US-08-232-246A-45 Sequence 45, Appl
128 21 84.0 107 4 US-08-557-050-5 Sequence 5, Appli
129 21 84.0 108 1 US-08-399-106A-1 Sequence 1, Appli
130 21 84.0 108 1 US-08-399-106A-2 Sequence 2, Appli
131 21 84.0 108 1 US-08-399-106A-3 Sequence 3, Appli
132 21 84.0 108 1 US-08-399-106A-8 Sequence 8, Appli
133 21 84.0 108 1 US-08-433-105A-1 Sequence 1, Appli
134 21 84.0 108 1 US-08-433-105A-2 Sequence 2, Appli
135 21 84.0 108 1 US-08-433-105A-3 Sequence 3, Appli
136 21 84.0 108 1 US-08-433-105A-8 Sequence 8, Appli
137 21 84.0 108 2 US-08-434-869A-1 Sequence 1, Appli
138 21 84.0 108 2 US-08-434-869A-2 Sequence 2, Appli
139 21 84.0 108 2 US-08-434-869A-3 Sequence 3, Appli
140 21 84.0 108 2 US-08-434-869A-8 Sequence 8, Appli
141 21 84.0 112 3 US-09-673-395A-526 Sequence 526, App
142 21 84.0 116 3 US-08-545-809A-143 Sequence 143, App
143 21 84.0 123 4 US-09-198-452A-868 Sequence 868, App
144 21 84.0 141 1 US-08-463-262A-4 Sequence 4, Appli
145 21 84.0 141 1 US-08-463-989-4 Sequence 4, Appli
146 21 84.0 141 3 US-09-003-574-4 Sequence 4, Appli
147 21 84.0 141 3 US-09-003-574-4 Sequence 4, Appli
148 21 84.0 141 4 US-09-864-541A-4 Sequence 4, Appli
149 21 84.0 145 1 US-08-463-262A-7 Sequence 7, Appli
150 21 84.0 145 1 US-08-463-989-7 Sequence 7, Appli

ALIGNMENTS

RESULT 1
US-09-107-532A-7091
; Sequence 7091, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660

COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 7091:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 144 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...144
; SEQUENCE DESCRIPTION: SEQ ID NO: 7091:
US-09-107-532A-7091

Query Match 100.0%; Score 25; DB 4; Length 144;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
DB 94 SSVMH 98

RESULT 2

US-09-934-551-2
; Sequence 2, Application US/09934551
; Patent No. 6461850
; GENERAL INFORMATION:
; APPLICANT: Li, Zhenya
; APPLICANT: BEASLEY, Ellen M.
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; FILE REFERENCE: CL000757
; CURRENT APPLICATION NUMBER: US/09/934,551
; CURRENT FILING DATE: 2001-08-23
; PRIOR APPLICATION NUMBER: 60/226,903
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 300
; TYPE: PRT
; ORGANISM: HUMAN
US-09-934-551-2

Query Match 100.0%; Score 25; DB 4; Length 300;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
DB 233 SSVMH 237

```

RESULT 3
US-10-224-414-2
; Sequence 2, Application US/10224414
; Patent No. 6638751
; GENERAL INFORMATION:
; APPLICANT: BEASLEY, Ellen M.
; APPLICANT: Li, Zhenya
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL000757 DIV
; CURRENT APPLICATION NUMBER: US/10/224,414
; CURRENT FILING DATE: 2002-08-21
; PRIOR APPLICATION NUMBER: 60/226,903
; PRIOR FILING DATE: 2001-08-23
; PRIOR APPLICATION NUMBER: 09/934,551
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 300
; TYPE: PRT
; ORGANISM: HUMAN
US-10-224-414-2

Query Match      100.0%; Score 25; DB 4; Length 300;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSVMH 5
Db 233 SSVMH 237

RESULT 4
US-09-489-039A-12736
; Sequence 12736, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 12736
; LENGTH: 406
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12736

Query Match      100.0%; Score 25; DB 4; Length 406;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSVMH 5
Db 273 SSVMH 277

RESULT 5
US-09-134-000C-4515
; Sequence 4515, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032

```

```

; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4515
; LENGTH: 517
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (477)..(500)
; OTHER INFORMATION: Amino acid 477 & 500 are Xaa wherein Xaa = any amino acid.
US-09-134-000C-4515

Query Match      100.0%; Score 25; DB 4; Length 517;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSVMH 5
Db 467 SSVMH 471

RESULT 6
5208144-37
; Patent No. 5208144
; APPLICANT: SMITH, JOHN A.; RAYCHOWDHURY, RAKTIMA; NILES, JOHN L.
; TITLE OF INVENTION: METHOD FOR DETECTION OF HUMAN DNA
; CONTAINING THE GENE ENCODING LOW DENSITY LIPOPROTEIN RECEPTOR
; NUMBER OF SEQUENCES: 42
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/396,697
; FILING DATE: 22-AUG-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 313,682
; FILING DATE: 22-FEB-1989
; APPLICATION NUMBER: 235,211
; FILING DATE: 23-AUG-1988
; SEQ ID NO: 37
; LENGTH: 726
5208144-37

Query Match      100.0%; Score 25; DB 6; Length 726;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSVMH 5
Db 82 SSVMH 86

RESULT 7
US-08-118-270-93
; Sequence 93, Application US/08118270
; Patent No. 5508384
; GENERAL INFORMATION:
; APPLICANT: Murphy, Randall B.
; APPLICANT: Schuster, David I.
; TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
; TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
; NUMBER OF SEQUENCES: 348
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/118,270
FILING DATE: 09-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/943,236
FILING DATE: 10-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Townsend, Kevin G.
REGISTRATION NUMBER: 34,033
REFERENCE/DOCKET NUMBER: MURPHY-2A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 93:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-118-270-93

Query Match 96.0%; Score 24; DB 1; Length 23;
Best Local Similarity 80.0%; Pred. No. 29;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
||:|
Db 13 SSIMH 17

RESULT 8
PCT-US93-08528-93
SEQUENCE 93, Application PC/TUS9308528
GENERAL INFORMATION:
APPLICANT: New York University
TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
NUMBER OF SEQUENCES: 348
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/08528
FILING DATE: 09-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/943,236
FILING DATE: 10-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Townsend, Kevin G.
REGISTRATION NUMBER: 34,033
REFERENCE/DOCKET NUMBER: MURPHY-2 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 93:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 amino acids
TYPE: amino acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US93-08528-93

Query Match 96.0%; Score 24; DB 5; Length 23;
Best Local Similarity 80.0%; Pred. No. 29;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
||:|
Db 13 SSIMH 17

RESULT 9
US-09-934-551-4
SEQUENCE 4, Application US/09934551
PATENT NO. 6461850
GENERAL INFORMATION:
APPLICANT: BEASLEY, Ellen M.
APPLICANT: Li, Zhenya
TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
TITLE OF INVENTION: USES THEREOF
FILE REFERENCE: CL000757
CURRENT APPLICATION NUMBER: US/09/934,551
CURRENT FILING DATE: 2001-08-23
PRIOR APPLICATION NUMBER: 60/226,903
PRIOR FILING DATE: 2001-08-23
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 268
TYPE: PRT
ORGANISM: HUMAN
US-09-934-551-4

Query Match 96.0%; Score 24; DB 4; Length 268;
Best Local Similarity 80.0%; Pred. No. 3.1e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
||:|
Db 219 SSIMH 223

RESULT 10
US-10-224-414-4
SEQUENCE 4, Application US/10224414
PATENT NO. 6638751
GENERAL INFORMATION:
APPLICANT: BEASLEY, Ellen M.
APPLICANT: Li, Zhenya
TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
TITLE OF INVENTION: USES THEREOF
FILE REFERENCE: CL000757 DIV
CURRENT APPLICATION NUMBER: US/10/224,414
CURRENT FILING DATE: 2002-08-21
PRIOR APPLICATION NUMBER: 60/226,903
PRIOR FILING DATE: 2001-08-23
PRIOR APPLICATION NUMBER: 09/934,551
PRIOR FILING DATE: 2001-08-23
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 268
TYPE: PRT
ORGANISM: HUMAN
US-10-224-414-4

Query Match 96.0%; Score 24; DB 4; Length 268;
Best Local Similarity 80.0%; Pred. No. 3.1e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
QY      1 SSVMH 5
Db      219 SSIMH 223

RESULT 11
US-09-134-001C-4615
; Sequence 4615, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4615
; LENGTH: 630
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4615

Query Match      96.0%; Score 24; DB 4; Length 630;
Best Local Similarity 80.0%; Pred. No. 7.1e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 SSVMH 5
Db      283 SSIMH 287

RESULT 12
US-09-441-502B-6
; Sequence 6, Application US/09441502B
; Patent No. 6455041
; GENERAL INFORMATION:
; APPLICANT: Dunbar, Bonita S.
; TITLE OF INVENTION: IMMUNOGENIC EPITOPES OF THE HUMAN ZONA PELLUCIDA PROTEIN
; FILE REFERENCE: 12231.2USU1
; CURRENT APPLICATION NUMBER: US/09/441,502B
; CURRENT FILING DATE: 1999-11-17
; NUMBER OF SEQ ID NOS: 104
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-441-502B-6

Query Match      88.0%; Score 22; DB 4; Length 19;
Best Local Similarity 80.0%; Pred. No. 67;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 SSVMH 5
Db      15 SSVLH 19

RESULT 13
US-09-441-502B-7
; Sequence 7, Application US/09441502B
; Patent No. 6455041
; GENERAL INFORMATION:
; APPLICANT: Dunbar, Bonita S.
; TITLE OF INVENTION: IMMUNOGENIC EPITOPES OF THE HUMAN ZONA PELLUCIDA PROTEIN
; FILE REFERENCE: 12231.2USU1
; CURRENT APPLICATION NUMBER: US/09/441,502B
```

```
; CURRENT FILING DATE: 1999-11-17
; NUMBER OF SEQ ID NOS: 104
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-441-502B-7

Query Match      88.0%; Score 22; DB 4; Length 19;
Best Local Similarity 80.0%; Pred. No. 67;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 SSVMH 5
Db      9 SSVLH 13

RESULT 14
US-08-118-270-232
; Sequence 232, Application US/08118270
; Patent No. 5508384
; GENERAL INFORMATION:
; APPLICANT: Murphy, Randall B.
; APPLICANT: Schuster, David I.
; TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
; TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
; NUMBER OF SEQUENCES: 348
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/118,270
; FILING DATE: 09-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/943,236
; FILING DATE: 10-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Townsend, Kevin G.
; REGISTRATION NUMBER: 34,033
; REFERENCE/DOCKET NUMBER: MURPHY=2A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 232:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-118-270-232

Query Match      88.0%; Score 22; DB 1; Length 35;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 SSVMH 5
Db      3 ASVMH 7

RESULT 15
```

PCT-US93-08528-232
; Sequence 232, Application PC/TUS9308528
; GENERAL INFORMATION:
; APPLICANT: New York University
; TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
; TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
; NUMBER OF SEQUENCES: 348
; CORRESPONDENCE ADDRESS:
; ADDRESSES: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/08528
; FILING DATE: 09-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/943,236
; FILING DATE: 10-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Townsend, Kevin G.
; REGISTRATION NUMBER: 34,033
; REFERENCE/DOCKET NUMBER: MURPHY=2 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 232:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US93-08528-232

Query Match 88.0%; Score 22; DB 5; Length 35;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
Db 3 ASVMH 7

RESULT 16
US-09-621-976-5485
; Sequence 5485, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 5485
; LENGTH: 71
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -56..-1
US-09-621-976-5485

Query Match 88.0%; Score 22; DB 4; Length 71;
Best Local Similarity 80.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
Db 38 SSVLH 42

RESULT 17

US-09-107-532A-7054
; Sequence 7054, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 7054:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 86 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (8) LOCATION 1...86
; SEQUENCE DESCRIPTION: SEQ ID NO: 7054:

US-09-107-532A-7054

Query Match 88.0%; Score 22; DB 4; Length 86;
Best Local Similarity 80.0%; Pred. No. 2.9e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
Db 71 NSVMH 75

RESULT 18
US-09-328-352-5968
; Sequence 5968, Application US/09328352

```

; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5968
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-5968

```

```

Query Match      88.0%; Score 22; DB 4; Length 201;
Best Local Similarity 80.0%; Pred. No. 6.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 SSVMH 5
       169 SIVMH 173

```

```

RESULT 19
US-09-328-352-4307
; Sequence 4307, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4307
; LENGTH: 220
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-4307

```

```

Query Match      88.0%; Score 22; DB 4; Length 220;
Best Local Similarity 80.0%; Pred. No. 7.2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 SSVMH 5
       27 SSVLH 31

```

```

RESULT 20
US-07-857-224B-27
; Sequence 27, Application US/07857224B
; Patent No. 5958784
; GENERAL INFORMATION:
; APPLICANT: Benner, Steven A.
; TITLE OF INVENTION: Predicting Folded Structures of Proteins
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Steven A. Benner
; STREET: Hadlaubstrasse 151
; CITY: Zurich
; STATE: none
; COUNTRY: Switzerland
; ZIP: (note: this is an international post code) CH-8092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette, 1.4 Mb storage
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.0
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/857,224B

```

```

; FILING DATE: 03/25/92
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA: none
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (International) 41 1 632 2830
; TELEFAX: (International) 41 1 262 2437
; TELEX: none
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 260
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: protein
; ORIGINAL SOURCE:
; ORGANISM: Saccharomyces cerevisiae
; FEATURE: Protein kinase; Table 8 Column 30
; PUBLICATION INFORMATION:
; AUTHORS:
; AUTHORS: Hanks, S. K.
; AUTHORS: Quinn, A. M.
; AUTHORS: Hunter, T.
; TITLE: The protein kinase family
; JOURNAL: Science
; VOLUME: 241
; PAGES: 42-52
; DATE: 1988
US-07-857-224B-27

```

```

Query Match      88.0%; Score 22; DB 2; Length 260;
Best Local Similarity 80.0%; Pred. No. 8.5e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 SSVMH 5
       208 SSVLH 212

```

```

RESULT 21
US-09-663-600A-190
; Sequence 190, Application US/09663600A
; Patent No. 6573068
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Duclert, Aymeric
; APPLICANT: Bouquelieret, Lydie
; TITLE OF INVENTION: EXTENDED CDNAS FOR SECRETED PROTEINS
; FILE REFERENCE: 31.US3.CIP
; CURRENT APPLICATION NUMBER: US/09/663,600A
; CURRENT FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/191,997
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: 60/066,677
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/069,957
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/074,121
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/081,563
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/096,116
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: 60/099,273
; PRIOR FILING DATE: 1998-09-04
; NUMBER OF SEQ ID NOS: 229
; SOFTWARE: Patent.pm
; SEQ ID NO 190
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -21...-1

```

US-09-663-600A-190

Query Match 88.0%; Score 22; DB 4; Length 267;
Best Local Similarity 80.0%; Pred. No. 8.8e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
:||||
Db 177 SSVLH 181

RESULT 22

US-09-489-039A-13210
Sequence 13210, Application US/09489039A

Patent No. 6610836
GENERAL INFORMATION:

APPLICANT: Gary Breton et. al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
FILE REFERENCE: 2709.2004001

CURRENT APPLICATION NUMBER: US/09/489,039A

PRIOR FILING DATE: 2000-01-27

PRIOR APPLICATION NUMBER: US 60/117,747

PRIOR FILING DATE: 1999-01-29

NUMBER OF SEQ ID NOS: 14342

SEQ ID NO 13210

LENGTH: 279

TYPE: PRT

ORGANISM: Klebsiella pneumoniae

US-09-489-039A-13210

Query Match 88.0%; Score 22; DB 4; Length 279;

Best Local Similarity 80.0%; Pred. No. 9.1e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
:||||
Db 161 ASVMH 165

RESULT 23

US-09-252-991A-19132
Sequence 19132, Application US/09252991A

Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252,991A

PRIOR FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788

PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190

PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 19132

LENGTH: 306

TYPE: PRT

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-19132

Query Match 88.0%; Score 22; DB 4; Length 306;

Best Local Similarity 80.0%; Pred. No. 1e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
:||||
Db 43 ASVMH 47

RESULT 24

US-09-134-000C-4441

Sequence 4441, Application US/09134000C

Patent No. 6617156

GENERAL INFORMATION:

APPLICANT: Lynn Doucette-Stamm et al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO

FILE REFERENCE: 032796-032

CURRENT APPLICATION NUMBER: US/09/134,000C

PRIOR FILING DATE: 1998-08-13

PRIOR APPLICATION NUMBER: US 60/055,778

NUMBER OF SEQ ID NOS: 6812

SOFTWARE: PatentIn version 3.1

SEQ ID NO 4441

LENGTH: 316

TYPE: PRT

ORGANISM: Enterococcus faecalis

US-09-134-000C-4441

Query Match 88.0%; Score 22; DB 4; Length 316;

Best Local Similarity 80.0%; Pred. No. 1e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
:||||
Db 67 SAVMH 71

RESULT 25

US-09-252-991A-21306

Sequence 21306, Application US/09252991A

Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252,991A

PRIOR FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788

PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190

PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 21306

LENGTH: 318

TYPE: PRT

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-21306

Query Match 88.0%; Score 22; DB 4; Length 318;

Best Local Similarity 80.0%; Pred. No. 1e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
:||||
Db 105 SAVMH 109

Search completed: September 24, 2004, 02:20:01

Job time : 11.4906 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 01:54:43 ; Search time 28.9623 Seconds
(without alignments)
55.513 Million cell updates/sec

Title: US-10-088-639a-2_COPY_158_162

Perfect score: 25
Sequence: 1 SSWMH 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1349238 seqs, 32158718 residues

Total number of hits satisfying chosen parameters: 1349238

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PTC_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/PTCUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	25	100.0	69	11	US-09-864-408A-5172
2	25	100.0	76	9	US-09-864-761-43019
3	25	100.0	96	12	US-10-276-774-2118
4	25	100.0	147	16	US-10-437-963-182859
5	25	100.0	252	14	US-10-145-586-37
6	25	100.0	253	11	US-09-981-566A-51
7	25	100.0	261	14	US-10-017-181-966
8	25	100.0	282	14	US-10-167-555-2
9	25	100.0	300	12	US-10-640-326-2
10	25	100.0	300	14	US-10-224-414-2
11	25	100.0	300	14	US-10-143-575-2
12	25	100.0	306	11	US-09-844-861A-24
13	25	100.0	314	9	US-09-886-055-85
14	25	100.0	314	9	US-09-886-055-89
15	25	100.0	314	10	US-09-804-291-85
					Sequence 5172, Ap
					Sequence 43019, A
					Sequence 2118, Ap
					Sequence 182859,
					Sequence 37, Appl
					Sequence 51, Appl
					Sequence 966, App
					Sequence 2, Appl
					Sequence 2, Appl
					Sequence 24, Appl
					Sequence 85, Appl
					Sequence 89, Appl
					Sequence 85, Appl

Sequence 89, Appl
Sequence 26, Appl
Sequence 4, Appl
Sequence 160, App
Sequence 162, App
Sequence 10, Appl
Sequence 878, App
Sequence 880, App
Sequence 162, App
Sequence 164, App
Sequence 756, App
Sequence 358, App
Sequence 758, App
Sequence 12, Appl
Sequence 20, Appl
Sequence 22, Appl
Sequence 56085, A
Sequence 73027, A
Sequence 75084, A
Sequence 75831, A
Sequence 60141, A
Sequence 2, Appl
Sequence 10822, A
Sequence 182856, A
Sequence 166420, A
Sequence 74, Appl
Sequence 37219, A
Sequence 41650, A
Sequence 256617, A
Sequence 136900, A
Sequence 49811, A
Sequence 35922, A
Sequence 48450, A
Sequence 164336, A
Sequence 5, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 159207, A
Sequence 4079, Ap
Sequence 153508, A
Sequence 64012, A
Sequence 2365, Ap
Sequence 12439, A
Sequence 205768, A
Sequence 15, Appl
Sequence 68667, A
Sequence 196, App
Sequence 79, Appl
Sequence 79, Appl
Sequence 34368, A
Sequence 162579, A
Sequence 3, Appl
Sequence 23731, A
Sequence 267001, A
Sequence 47, Appl
Sequence 47, Appl
Sequence 762, App
Sequence 762, App
Sequence 762, App
Sequence 167025, A
Sequence 220384, A
Sequence 214496, A
Sequence 238615, A
Sequence 2100, Ap
Sequence 210914, A
Sequence 58681, A
Sequence 208508, A
Sequence 4981, Ap
Sequence 149134, A
Sequence 216314, A
Sequence 85, Appl
Sequence 272562, A

```
89 22 88.0 84 14 US-10-072-349-85
90 22 88.0 85 16 US-10-437-963-199497
91 22 88.0 86 11 US-09-864-408A-7902
92 22 88.0 87 16 US-10-437-963-174887
93 22 88.0 88 16 US-10-437-963-182752
94 22 88.0 89 16 US-10-437-963-183315
95 22 88.0 90 10 US-09-764-891-4091
96 22 88.0 93 12 US-10-424-599-196992
97 22 88.0 97 16 US-10-437-963-135184
98 22 88.0 98 12 US-09-978-360A-524
99 22 88.0 98 15 US-10-291-265-749
100 22 88.0 101 12 US-10-424-599-248946
101 22 88.0 102 9 US-09-925-299-783
102 22 88.0 102 10 US-09-925-299-783
103 22 88.0 103 12 US-10-425-114-39573
104 22 88.0 104 15 US-10-264-049-4113
105 22 88.0 107 12 US-10-424-599-220888
106 22 88.0 108 9 US-09-764-868-745
107 22 88.0 110 12 US-10-276-774-1431
108 22 88.0 110 16 US-10-767-701-59432
109 22 88.0 111 12 US-10-425-114-45136
110 22 88.0 116 16 US-10-437-963-120695
111 22 88.0 118 12 US-10-424-599-268571
112 22 88.0 124 16 US-10-437-963-116901
113 22 88.0 124 16 US-10-437-963-203892
114 22 88.0 124 16 US-10-767-701-61619
115 22 88.0 141 16 US-10-437-963-174244
116 22 88.0 149 12 US-10-424-599-284441
117 22 88.0 150 12 US-10-424-599-159834
118 22 88.0 150 14 US-10-002-631C-8
119 22 88.0 155 14 US-10-050-704-175
120 22 88.0 155 16 US-10-798-512-175
121 22 88.0 156 12 US-10-424-599-224708
122 22 88.0 157 16 US-10-767-701-36163
123 22 88.0 160 16 US-10-437-963-129045
124 22 88.0 164 12 US-10-425-114-40783
125 22 88.0 165 12 US-10-425-114-44317
126 22 88.0 167 16 US-10-367-094-132
127 22 88.0 169 15 US-10-108-260A-3550
128 22 88.0 170 16 US-10-437-963-108761
129 22 88.0 173 16 US-10-767-701-43865
130 22 88.0 173 16 US-10-767-701-53022
131 22 88.0 177 12 US-10-282-122A-54047
132 22 88.0 178 12 US-10-424-599-156228
133 22 88.0 181 10 US-09-791-932-100
134 22 88.0 184 12 US-10-424-599-163657
135 22 88.0 191 12 US-10-425-114-61750
136 22 88.0 191 12 US-10-425-114-65124
137 22 88.0 191 14 US-10-101-464A-625
138 22 88.0 199 12 US-10-424-599-192997
139 22 88.0 201 15 US-10-369-493-2524
140 22 88.0 202 12 US-10-424-599-250163
141 22 88.0 204 12 US-10-343-650A-174
142 22 88.0 204 14 US-10-017-161-2170
143 22 88.0 204 15 US-10-120-801-58
144 22 88.0 204 15 US-10-292-798-1816
145 22 88.0 205 15 US-10-369-493-20215
146 22 88.0 209 14 US-10-002-631C-254
147 22 88.0 211 16 US-10-437-963-187108
148 22 88.0 213 12 US-10-335-977-6562
149 22 88.0 216 9 US-09-811-284-181
150 22 88.0 218 16 US-10-767-701-31924
```

ALIGNMENTS

```
RESULT 1
US-09-864-408A-5172
; Sequence 5172, Application US/09864408A
; Publication No. US20040009474A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
```

```
; APPLICANT: Shimkets, Richard A.
; TITLE OF INVENTION: No. US20040009474A1el Human Polynucleotides and Polypeptides Encor
; FILE REFERENCE: 21402-012
; CURRENT APPLICATION NUMBER: US/09/864,408A
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/206,690
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 9068
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5172
; LENGTH: 69
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-864-408A-5172

Query Match 100.0%; Score 25; DB 11; Length 69;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
DB 23 SSVMH 27

RESULT 2
US-09-864-761-43019
; Sequence 43019, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aemica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
```

; PRIOR FILING DATE: 2001-01-29
 ; NUMBER OF SEQ ID NOS: 49117
 ; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
 ; SEQ ID NO 43019
 ; LENGTH: 76
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: MAP TO AC009758.1
 ; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
 ; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.92
 ; OTHER INFORMATION: SWISSPROT HIT: Q95156, EVALUATE 6.00e-05
 ; OTHER INFORMATION: EST HUMAN HIT: A1138213.1, EVALUATE 4.00e-25
 ; OTHER INFORMATION: EST_HUMAN HIT: A1138213.1, EVALUATE 4.00e-25
 ; NAME/KEY: unsure
 ; LOCATION: 13
 ; OTHER INFORMATION: Clone ID: PAT_MRT4530_80004C.1.pap
 US-10-437-963-182859

Query Match 100.0%; Score 25; DB 9; Length 76;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
 DB 35 SSVMH 39

RESULT 3
 US-10-276-774-2118
 ; Sequence 2118, Application US/10276774
 ; Publication No. US20040053245A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hyseq, Inc.
 ; APPLICANT: Tang, Y. Tom et al
 ; TITLE OF INVENTION: NO. US20040053245A1el Nucleic Acids and Polypeptides
 ; FILE REFERENCE: 21272-030
 ; CURRENT APPLICATION NUMBER: US/10/276,774
 ; CURRENT FILING DATE: 2002-11-18
 ; PRIOR FILING DATE: 2002-11-18
 ; PRIOR FILING DATE: 2000-04-27
 ; PRIOR APPLICATION NUMBER: 09/560,875
 ; PRIOR FILING DATE: 2000-04-27
 ; PRIOR APPLICATION NUMBER: 09/496,914
 ; PRIOR FILING DATE: 2000-02-03
 ; NUMBER OF SEQ ID NOS: 2700
 ; SOFTWARE: Custom
 ; SEQ ID NO 2118
 ; LENGTH: 96
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-276-774-2118

Query Match 100.0%; Score 25; DB 12; Length 96;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
 DB 39 SSVMH 43

RESULT 4
 US-10-437-963-182859
 ; Sequence 182859, Application US/10437963
 ; Publication No. US20040123343A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Wu, Wei
 ; APPLICANT: Boukharov, Andrey A.
 ; APPLICANT: Barbazuk, Brad
 ; APPLICANT: Li, Ping
 ; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53221)B
 ; CURRENT APPLICATION NUMBER: US/10/437,963
 ; CURRENT FILING DATE: 2003-05-14
 ; NUMBER OF SEQ ID NOS: 204966
 ; SEQ ID NO 182859
 ; LENGTH: 147
 ; TYPE: PRT
 ; ORGANISM: Oryza sativa
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (1)..(147)
 ; OTHER INFORMATION: unsure at all xaa locations
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT4530_80004C.1.pap
 US-10-437-963-182859

Query Match 100.0%; Score 25; DB 16; Length 147;
 Best Local Similarity 100.0%; Pred. No. 3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
 DB 60 SSVMH 64

RESULT 5
 US-10-145-586-37
 ; Sequence 37, Application US/10145586
 ; Publication No. US20030138890A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Alexandra Glucksmann, Maria
 ; APPLICANT: Silos-Santiago, Inmaculada
 ; APPLICANT: M. Galvin, Katherine
 ; APPLICANT: Weich, Nadine
 ; APPLICANT: Curtis, Rory A.J.
 ; APPLICANT: Bandaru, Rajasekhar
 ; APPLICANT: Kapeller-Libermann, Rosana
 ; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTOR FAMILY MEMBERS,
 ; TITLE OF INVENTION: HUMAN THIOREDOXIN FAMILY MEMBERS, HUMAN LEUCINE-RICH
 ; FILE REFERENCE: 10448-188001
 ; CURRENT APPLICATION NUMBER: US/10/145,586
 ; CURRENT FILING DATE: 2002-05-14
 ; Prior Application removed - See File Wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 95
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 37
 ; LENGTH: 252
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-145-586-37

Query Match 100.0%; Score 25; DB 14; Length 252;
 Best Local Similarity 100.0%; Pred. No. 5.2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
 DB 218 SSVMH 222

RESULT 6
 US-09-981-566A-51
 ; Sequence 51, Application US/09981566A
 ; Publication No. US20040005656A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kekuda et al.
 ; TITLE OF INVENTION: No. US20040005656A1el GPCR-like Proteins and Nucleic Acids Encod
 ; TITLE OF INVENTION: Same
 ; FILE REFERENCE: 21402-163
 ; CURRENT APPLICATION NUMBER: US/09/981,566A
 ; CURRENT FILING DATE: 2001-10-16

PRIOR APPLICATION NUMBER: 60/240,704
PRIOR FILING DATE: 2000-10-16
PRIOR APPLICATION NUMBER: 60/262,159
PRIOR FILING DATE: 2001-01-17
PRIOR APPLICATION NUMBER: 60/263,340
PRIOR FILING DATE: 2001-01-22
PRIOR APPLICATION NUMBER: 60/264,118
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 60/308,203
PRIOR FILING DATE: 2001-07-27
PRIOR APPLICATION NUMBER: 60/243,497
PRIOR FILING DATE: 2000-10-26
PRIOR APPLICATION NUMBER: 60/244,542
PRIOR FILING DATE: 2000-10-31
PRIOR APPLICATION NUMBER: 60/269,031
PRIOR FILING DATE: 2001-02-15
PRIOR APPLICATION NUMBER: 60/245,484
PRIOR FILING DATE: 2000-11-03
PRIOR APPLICATION NUMBER: 60/255,017
PRIOR FILING DATE: 2000-12-12
PRIOR APPLICATION NUMBER: 60/263,216
PRIOR FILING DATE: 2001-01-22
PRIOR APPLICATION NUMBER: 60/268,225
PRIOR FILING DATE: 2001-02-12
NUMBER OF SEQ ID NOS: 209
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 51
LENGTH: 253
TYPE: PRT
ORGANISM: Homo sapiens
US-09-981-566A-51

Query Match 100.0%; Score 25; DB 11; Length 253;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
Db 199 SSVMH 203

RESULT 7
US-10-017-161-966
Sequence 966, Application US/10017161
Publication No. US2003014368A1
GENERAL INFORMATION:
APPLICANT: SUWA, MAKIKO
APPLICANT: ASAI, KIYOSHI
APPLICANT: AKIYAMA, YUTAKA
APPLICANT: ABURATANI, HIROYUKI
TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
FILE REFERENCE: 084335/0152
CURRENT APPLICATION NUMBER: US/10/017,161
CURRENT FILING DATE: 2002-12-18
PRIOR APPLICATION NUMBER: JP 2001/246789
PRIOR FILING DATE: 2001-06-18
NUMBER OF SEQ ID NOS: 2430
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 966
LENGTH: 261
TYPE: PRT
ORGANISM: Homo sapiens
US-10-017-161-966

Query Match 100.0%; Score 25; DB 14; Length 261;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
Db 204 SSVMH 208

RESULT 8

US-10-167-555-2
Sequence 2, Application US/10167555
Publication No. US20030022212A1
GENERAL INFORMATION:
APPLICANT: Curtis, Rory A. J.
TITLE OF INVENTION: 65649, A Human Metalloprotease Family Member and Uses
FILE REFERENCE: MPI01-089P1RM
CURRENT APPLICATION NUMBER: US/10/167,555
CURRENT FILING DATE: 2002-06-12
PRIOR APPLICATION NUMBER: 60/297,938
PRIOR FILING DATE: 2001-06-13
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 282
TYPE: PRT
ORGANISM: Homo sapiens
US-10-167-555-2

Query Match 100.0%; Score 25; DB 14; Length 282;
Best Local Similarity 100.0%; Pred. No. 5.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
Db 233 SSVMH 237

RESULT 9

US-10-640-326-2
Sequence 2, Application US/10640326
Publication No. US2004003896A1
GENERAL INFORMATION:
APPLICANT: BEASLEY, Ellen M.
APPLICANT: LI, Zhenya
TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
TITLE OF INVENTION: USES THEREOF
FILE REFERENCE: CLO00757 DIV2
CURRENT APPLICATION NUMBER: US/10/640,326
CURRENT FILING DATE: 2003-08-25
PRIOR APPLICATION NUMBER: 60/226,903
PRIOR FILING DATE: 2000-08-23
PRIOR APPLICATION NUMBER: 09/934,551
PRIOR FILING DATE: 2001-8-23
PRIOR APPLICATION NUMBER: 10/224,414
PRIOR FILING DATE: 2002-08-21
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 300
TYPE: PRT
ORGANISM: HUMAN
US-10-640-326-2

Query Match 100.0%; Score 25; DB 12; Length 300;
Best Local Similarity 100.0%; Pred. No. 6.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
Db 233 SSVMH 237

RESULT 10

US-10-224-414-2
Sequence 2, Application US/10224414
Publication No. US20030036167A1
GENERAL INFORMATION:
APPLICANT: BEASLEY, Ellen M.
APPLICANT: LI, Zhenya

```

; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; FILE REFERENCE: CL000757 DIV
; CURRENT APPLICATION NUMBER: US/10/224,414
; CURRENT FILING DATE: 2002-08-21
; PRIOR APPLICATION NUMBER: 60/226,903
; PRIOR FILING DATE: 2001-08-23
; PRIOR APPLICATION NUMBER: 09/934,551
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 300
; TYPE: PRT
; ORGANISM: HUMAN
US-10-224-414-2

Query Match      100.0%; Score 25; DB 14; Length 300;
Best Local Similarity 100.0%; Pred. No. 6.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SSVMH 5
        |||||
DB      233 SSVMH 237

RESULT 11
US-10-143-575-2
; Sequence 2, Application US/10143575
; Publication No. US20030166072A1
; GENERAL INFORMATION:
; APPLICANT: ZHONG, Jenny M. et al.
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; FILE REFERENCE: CL001240
; CURRENT APPLICATION NUMBER: US/10/143,575
; CURRENT FILING DATE: 2002-05-13
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-143-575-2

Query Match      100.0%; Score 25; DB 14; Length 300;
Best Local Similarity 100.0%; Pred. No. 6.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SSVMH 5
        |||||
DB      233 SSVMH 237

RESULT 12
US-09-844-861A-24
; Sequence 24, Application US/09844861A
; Publication No. US20030216304A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Mishra, Vishnu
; APPLICANT: Spytek, Kimberly
; APPLICANT: Burgess, Catherine
; APPLICANT: Lepley, Denise
; APPLICANT: Grosse, William
; APPLICANT: Szekeres, Edward
; APPLICANT: Alsobrook, John
; APPLICANT: Gangolli, Esha
; APPLICANT: Casman, Stacie
; APPLICANT: MacDougall, John
; APPLICANT: Smithson, Glennda

```

```

; TITLE OF INVENTION: NO. US20030216304A1e1 Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-789 US
; CURRENT APPLICATION NUMBER: US/09/844,861A
; CURRENT FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: 60/199,947
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/199,960
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/225,226
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/256,399
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/256,524
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/258,159
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/258,511
; PRIOR FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: 60/258,828
; PRIOR FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: 60/259,659
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/275,604
; PRIOR FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 113
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 24
; LENGTH: 306
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-844-861A-24

Query Match      100.0%; Score 25; DB 11; Length 306;
Best Local Similarity 100.0%; Pred. No. 6.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SSVMH 5
        |||||
DB      249 SSVMH 253

RESULT 13
US-09-886-055-85
; Sequence 85, Application US/09886055
; Patent No. US20020132273A1
; GENERAL INFORMATION:
; APPLICANT: STRYER, LUBERT
; APPLICANT: ZOZULYA, SERGEY
; TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND
; TITLE OF INVENTION: BIOSENSORS OF CHEMICAL SENSANTS
; FILE REFERENCE: 078003-0277150
; CURRENT APPLICATION NUMBER: US/09/886,055
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/213,812
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 522
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 85
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-886-055-85

Query Match      100.0%; Score 25; DB 9; Length 314;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SSVMH 5
        |||||
DB      260 SSVMH 264

RESULT 14

```

```

US-09-886-055-89
; Sequence 89, Application US/09886055
; Patent No. US20020132273A1
; GENERAL INFORMATION:
; APPLICANT: STRYER, LUBERT
; APPLICANT: ZOZULYA, SERGEY
; TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND
; TITLE OF INVENTION: BIOSENSORS OF CHEMICAL SENSANTS
; FILE REFERENCE: 078003-0277150
; CURRENT APPLICATION NUMBER: US/09/886,055
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/213,812
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 522
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 89
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-886-055-89

Query Match      100.0%; Score 25; DB 9; Length 314;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SSVMH 5
Db      260 SSVMH 264

RESULT 15
US-09-804-291-85
; Sequence 85, Application US/09804291
; Publication No. US20030088059A1
; GENERAL INFORMATION:
; APPLICANT: ZOZULYA, SERGEY
; TITLE OF INVENTION: HUMAN OLFACTORY RECEPTORS AND GENES ENCODING SAME
; FILE REFERENCE: P 0278005
; CURRENT APPLICATION NUMBER: US/09/804,291
; CURRENT FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: 60/188,914
; PRIOR FILING DATE: 2000-03-13
; PRIOR APPLICATION NUMBER: 60/192,033
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/198,474
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/199,335
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/213,849
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/226,534
; PRIOR FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: 60/230,732
; PRIOR FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: 60/266,862
; PRIOR FILING DATE: 2001-02-07
; NUMBER OF SEQ ID NOS: 529
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 85
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-804-291-85

Query Match      100.0%; Score 25; DB 10; Length 314;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SSVMH 5
Db      260 SSVMH 264

RESULT 16
US-09-804-291-89
; Sequence 89, Application US/09804291
; Publication No. US20030088059A1
; GENERAL INFORMATION:
; APPLICANT: ZOZULYA, SERGEY
; TITLE OF INVENTION: HUMAN OLFACTORY RECEPTORS AND GENES ENCODING SAME
; FILE REFERENCE: P 0278005
; CURRENT APPLICATION NUMBER: US/09/804,291
; CURRENT FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: 60/188,914
; PRIOR FILING DATE: 2000-03-13
; PRIOR APPLICATION NUMBER: 60/192,033
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/198,474
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/199,335
; PRIOR FILING DATE: 2000-04-24
; PRIOR APPLICATION NUMBER: 60/207,702
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/213,849
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/226,534
; PRIOR FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: 60/230,732
; PRIOR FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: 60/266,862
; PRIOR FILING DATE: 2001-02-07
; NUMBER OF SEQ ID NOS: 529
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 89
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-804-291-89

Query Match      100.0%; Score 25; DB 10; Length 314;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SSVMH 5
Db      260 SSVMH 264

RESULT 17
US-09-844-861A-26
; Sequence 26, Application US/09844861A
; Publication No. US20030216304A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Mishra, Vishnu
; APPLICANT: Spytek, Kimberly
; APPLICANT: Burgess, Catherine
; APPLICANT: Lepley, Denise
; APPLICANT: Grosse, William
; APPLICANT: Szekeres, Edward
; APPLICANT: Alsobrook, John
; APPLICANT: Gangolli, Esha
; APPLICANT: Casman, Stacie
; APPLICANT: MacDougall, John
; APPLICANT: Smithson, Glennda
; TITLE OF INVENTION: No. US20030216304A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-789 US
; CURRENT APPLICATION NUMBER: US/09/844,861A
; CURRENT FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: 60/199,947
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/199,960
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/225,226

```

```
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/256,399
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/256,524
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/258,159
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/258,511
; PRIOR FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: 60/258,828
; PRIOR FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: 60/259,659
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/275,604
; PRIOR FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 113
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-844-861A-26

Query Match      100.0%; Score 25; DB 11; Length 314;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SSMVH 5
Db      260 SSMVH 264

RESULT 18
US-10-182-822A-4
; Sequence 4, Application US/10182822A
; Publication No. US20030211493A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.; BAUGHN, Mariah R.
; APPLICANT: AU-YOUNG, Janice; YUE, Henry
; TITLE OF INVENTION: G-PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: PI-0032 USN
; CURRENT APPLICATION NUMBER: US/10/182,822A
; CURRENT FILING DATE: 2001-02-01
; PRIOR APPLICATION NUMBER: PCT/US 01/03455
; PRIOR FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: US 60/180,093
; PRIOR FILING DATE: 2000-02-02
; PRIOR APPLICATION NUMBER: US 60/182,045
; PRIOR FILING DATE: 2000-02-11
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PERL Program
; SEQ ID NO 4
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 7472036CD1
US-10-182-822A-4

Query Match      100.0%; Score 25; DB 12; Length 314;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SSMVH 5
Db      260 SSMVH 264

RESULT 19
US-10-343-650A-160
; Sequence 160, Application US/10343650A
; Publication No. US20040067499A1
```

```
; GENERAL INFORMATION:
; APPLICANT: HAGA, TATSUYA
; TITLE OF INVENTION: NOVEL G-PROTEIN COUPLED RECEPTOR
; FILE REFERENCE: 31671-186347 US/10/343,650A
; CURRENT APPLICATION NUMBER: US/10/343,650A
; CURRENT FILING DATE: 2003-07-21
; PRIOR APPLICATION NUMBER: JP 2000/237818
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: JP 2001/34434
; PRIOR FILING DATE: 2001-02-13
; NUMBER OF SEQ ID NOS: 694
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 160
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-343-650A-160

Query Match      100.0%; Score 25; DB 12; Length 314;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SSMVH 5
Db      260 SSMVH 264

RESULT 20
US-10-343-650A-162
; Sequence 162, Application US/10343650A
; Publication No. US20040067499A1
; GENERAL INFORMATION:
; APPLICANT: HAGA, TATSUYA
; TITLE OF INVENTION: NOVEL G-PROTEIN COUPLED RECEPTOR
; FILE REFERENCE: 31671-186347
; CURRENT APPLICATION NUMBER: US/10/343,650A
; CURRENT FILING DATE: 2003-07-21
; PRIOR APPLICATION NUMBER: JP 2000/237818
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: JP 2001/34434
; PRIOR FILING DATE: 2001-02-13
; NUMBER OF SEQ ID NOS: 694
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 162
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-343-650A-162

Query Match      100.0%; Score 25; DB 12; Length 314;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SSMVH 5
Db      260 SSMVH 264

RESULT 21
US-10-032-106-10
; Sequence 10, Application US/10032106
; Publication No. US20020155473A1
; GENERAL INFORMATION:
; APPLICANT: Feng, Zaoyuan
; APPLICANT: Yu, Zailin
; APPLICANT: Wiley, Andrew
; APPLICANT: Hu, Qianjin
; TITLE OF INVENTION: METHODS FOR IDENTIFYING G-PROTEIN
; FILE REFERENCE: 433112000700
; CURRENT APPLICATION NUMBER: US/10/032,106
; CURRENT FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: US 60/258,070
```

```
; PRIOR FILING DATE: 2000-12-20
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-032-106-10

Query Match      100.0%; Score 25; DB 13; Length 314;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SSVMH 5
DB      260 SSVMH 264
      |||||

RESULT 22
US-10-017-161-878
; Sequence 878, Application US/10017161
; Publication No. US20030143668A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 084335/0152
; CURRENT APPLICATION NUMBER: US/10/017,161
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP 2001/246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2430
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 878
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-017-161-878

Query Match      100.0%; Score 25; DB 14; Length 314;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SSVMH 5
DB      260 SSVMH 264
      |||||

RESULT 23
US-10-017-161-880
; Sequence 880, Application US/10017161
; Publication No. US20030143668A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 084335/0152
; CURRENT APPLICATION NUMBER: US/10/017,161
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP 2001/246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2430
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 880
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-017-161-880

Query Match      100.0%; Score 25; DB 13; Length 314;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SSVMH 5
DB      260 SSVMH 264
      |||||

RESULT 24
US-10-387-629-162
; Sequence 162, Application US/10387629
; Publication No. US20030221205A1
; GENERAL INFORMATION:
; APPLICANT: Veithen, Alex
; APPLICANT: ChemCom S.A.
; TITLE OF INVENTION: Olfactory and Pheromones G-Protein coupled Receptors
; FILE REFERENCE: 9409/2192
; CURRENT APPLICATION NUMBER: US/10/387,629
; CURRENT FILING DATE: 2003-03-13
; NUMBER OF SEQ ID NOS: 254
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 162
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-387-629-162

Query Match      100.0%; Score 25; DB 15; Length 314;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SSVMH 5
DB      260 SSVMH 264
      |||||

RESULT 25
US-10-387-629-164
; Sequence 164, Application US/10387629
; Publication No. US20030221205A1
; GENERAL INFORMATION:
; APPLICANT: Veithen, Alex
; APPLICANT: ChemCom S.A.
; TITLE OF INVENTION: Olfactory and Pheromones G-Protein coupled Receptors
; FILE REFERENCE: 9409/2192
; CURRENT APPLICATION NUMBER: US/10/387,629
; CURRENT FILING DATE: 2003-03-13
; NUMBER OF SEQ ID NOS: 254
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 164
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-387-629-164

Query Match      100.0%; Score 25; DB 15; Length 314;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SSVMH 5
DB      260 SSVMH 264
      |||||

Search completed: September 24, 2004, 02:25:13
Job time : 32.9623 secs
```

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	25	100.0	112	2	AF30502	Ig heavy chain V r conserved hypot het metalloproteinase molybdenum transpo homeotic protein z ubiquinol-cytochro ubiquinol-cytochro Rhusus-like protei Rhusus-like protei Rhusus-like protei related to DRPLA p transcription fact gp130 protein prec DNA-directed RNA p hypot hetical prote hypot hetical prote hypot hetical prote hypot hetical prote high choriotyic h high choriotyic h hypot hetical prote ABC transporter AT probable transmemb hypot hetical prote A11i protein - var A10i protein - var hypot hetical prote hypot hetical prote
2	25	100.0	221	2	AF2871	
3	25	100.0	310	2	S41055	
4	25	100.0	352	2	AE0595	
5	25	100.0	353	2	A43697	
6	25	100.0	385	1	S01511	
7	25	100.0	396	2	S54999	
8	25	100.0	417	2	I37004	
9	25	100.0	417	2	I37075	
10	25	100.0	417	2	I37076	
11	25	100.0	417	2	I37003	
12	25	100.0	538	2	T49849	
13	25	100.0	1065	2	A57410	
14	25	100.0	4660	2	T42737	
15	24	96.0	90	2	A82044	
16	24	96.0	186	2	H86154	
17	24	96.0	199	2	H86154	
18	24	96.0	247	2	S57548	
19	24	96.0	258	2	T22233	
20	24	96.0	270	2	B48826	
21	24	96.0	279	2	C48826	
22	24	96.0	330	2	B69416	
23	24	96.0	496	2	C87194	
24	24	96.0	646	2	AD3409	
25	24	96.0	767	2	T30018	
26	24	96.0	892	2	H72164	
27	24	96.0	892	2	B36849	
28	24	96.0	892	2	T28552	
29	22	88.0	51	2	T07326	

103 22 88.0 472 1 YKJUM
 104 22 88.0 472 2 T41684
 105 22 88.0 474 1 S75568
 106 22 88.0 475 2 JX0344
 107 22 88.0 481 2 J01147
 108 22 88.0 488 2 T32149
 109 22 88.0 490 1 OARBP2
 110 22 88.0 490 2 A5954
 111 22 88.0 490 2 I48162
 112 22 88.0 490 2 I48163
 113 22 88.0 490 2 I48189
 114 22 88.0 490 2 I49610
 115 22 88.0 492 2 D35114
 116 22 88.0 492 2 T28025
 117 22 88.0 493 2 E35115
 118 22 88.0 493 2 T07374
 119 22 88.0 498 2 S78184
 120 22 88.0 500 2 G71836
 121 22 88.0 500 2 B64680
 122 22 88.0 502 2 A41361
 123 22 88.0 505 2 A39128
 124 22 88.0 505 2 C86180
 125 22 88.0 510 2 T04145
 126 22 88.0 520 2 C70776
 127 22 88.0 540 2 S70403
 128 22 88.0 544 2 S41094
 129 22 88.0 544 2 S41095
 130 22 88.0 544 2 S41096
 131 22 88.0 548 2 J40758
 132 22 88.0 548 2 T47548
 133 22 88.0 548 2 T40508
 134 22 88.0 551 2 T00932
 135 22 88.0 556 2 S67097
 136 22 88.0 562 2 T05758
 137 22 88.0 563 1 P04093
 138 22 88.0 565 2 S67298
 139 22 88.0 571 2 C84038
 140 22 88.0 576 2 T12005
 141 22 88.0 584 2 G71676
 142 22 88.0 597 2 B6251
 143 22 88.0 626 2 H82133
 144 22 88.0 649 2 I64847
 145 22 88.0 649 2 I38204
 146 22 88.0 649 2 E84614
 147 22 88.0 650 2 A54065
 148 22 88.0 654 2 F71298
 149 22 88.0 664 2 G89894
 150 22 88.0 668 1 A42908

ALIGNMENTS

RESULT 1
 A30502
 Ig heavy chain V region (D44) - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 03-Nov-1988 #sequence_revision 03-Nov-1988 #text_change 21-Jan-2000
 C:Accession: A30502
 R:Eilat, D.; Webster, D.M.; Rees, A.R.
 J. Immunol. 141, 1745-1753, 1988
 A:Title: V region sequences of anti-DNA and anti-RNA autoantibodies from NZB/NZW F-1 mice
 A:Reference number: A30502; MUID:88315787; PMID:2457627
 A:Accession: A30502
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-112 <EIL>
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotrimer; immunoglobulin
 F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 25; DB 2; Length 112;
 Best Local Similarity 100.0%; Pred. No. 17;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SSVMH 5
 DB 31 SSVMH 35

RESULT 2

AF2871
 conserved hypothetical protein Atu2400 [imported] - Agrobacterium tumefaciens (strain C58)
 C:Species: Agrobacterium tumefaciens
 C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
 C:Accession: AF2871
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.;
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kuryavin, T.; Levy, R.; Li, M.; McClellan,
 ; Karp, P.; Romero, P.; Grant, P.; Zhang, S.
 Science 294, 2317-2323, 2001
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, E.
 ster, E.W.
 A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
 A:Reference number: AB2577; MUID:21608550; PMID:11743193
 A:Accession: AF2871
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-221 <KUR>
 A:Cross-references: GB:AE008688; PIDN:AAL43388.1; PID:gl7740886; GSPDB:GN00186
 A:Experimental source: strain C58 (Dupont)
 C:Genetics:
 A:Gene: Atu2400
 A:Map position: circular chromosome

Query Match 100.0%; Score 25; DB 2; Length 221;
 Best Local Similarity 100.0%; Pred. No. 35;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
 DB 115 SSVMH 119

RESULT 3

S41055
 metalloproteinase (EC 3.4.24.-) - Japanese quail
 C:Species: Coturnix coturnix japonica (Japanese quail)
 C:Date: 09-Jun-1994 #sequence_revision 10-Nov-1995 #text_change 28-Jul-2000
 C:Accession: S41055
 R:Elaroussi, M.A.; Deluca, H.F.
 Biochim. Biophys. Acta 1217, 1-8, 1994
 A:Title: A new member to the astacin family of metalloendopeptidases: a novel 1,25-dihydroxy-
 A:Reference number: S41055; MUID:94114563; PMID:8286408
 A:Accession: S41055
 A:Molecule type: mRNA
 A:Residues: 1-310 <ELA>
 A:Cross-references: GB:U12642; GB:S68488; NID:g530065; PIDN:AAA20842.1; PID:g530066
 A:Note: the authors translated the codon AAC for residue 59 as Trp
 C:Superfamily: Japanese quail metalloproteinase; astacin homology; C1r/C1s repeat homolog
 C:Keywords: hydrolase; metalloproteinase; zinc
 F:3-187/Domain: astacin homology <AST>
 F:188-297/Domain: C1r/C1s repeat homology <C1R2>
 F:84,88,94,141/Binding site: zinc (His, His, His, Tyr) #status predicted
 F:85/Active site: Glu #status predicted

Query Match 100.0%; Score 25; DB 2; Length 310;
 Best Local Similarity 100.0%; Pred. No. 52;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
 DB 136 SSVMH 140

RESULT 4

AE0595

molybdenum transport ATP-binding protein ModC STY0816 [imported] - Salmonella enterica
 C:Species: Salmonella enterica subsp. enterica serovar Typhi
 A:Note: This species has also been called Salmonella typhi
 C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
 C:Accession: AE0595
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
 th, T.; Connerton, F.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
 S.; Moule, S.; O'Gaora, P.
 Nature 413, 848-852, 2001
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
 A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
 A:Reference number: AB0502; MUID:21534947; PMID:11677608
 A:Accession: AE0595
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-352 <PAR>
 A:Cross-references: GB:AL513382; PID:CAD05231.1; PID:g16502001; GSPDB:GN00176
 C:Genetics:
 A:Gene: STY0816
 C:Superfamily: molybdenum transport protein modC; ATP-binding cassette homology

Query Match 100.0%; Score 25; DB 2; Length 352;
 Best Local Similarity 100.0%; Pred. No. 60;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSVMH 5
 Db 219 SSVMH 223
 |||||

RESULT 5
 A43697
 homeotic protein zen1 - fruit fly (Drosophila melanogaster)
 C:Species: Drosophila melanogaster
 C:Date: 03-Mar-1993 #sequence_revision 03-Mar-1993 #text_change 24-Sep-1999
 C:Accession: A43697
 R:Rushlow, C.; Doyle, H.; Hoey, T.; Levine, M.
 Genes Dev. 1, 1268-1279, 1987
 A:Title: Molecular characterization of the zerknullt region of the antennapedia gene co
 A:Reference number: A43697; MUID:88112803; PMID:2892759
 A:Accession: A43697
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-353 <RUS>
 A:Cross-references: GB:X68347; NID:g807980; PIDN:CAA48416.1; PID:g807981
 C:Genetics:
 A:Gene: FlyBase:zen
 A:Cross-references: FlyBase:FBgn0004053
 C:Superfamily: unassigned homeobox proteins; homeobox homology
 C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
 F:91-147/Domain: homeobox homology <HOK>

Query Match 100.0%; Score 25; DB 2; Length 353;
 Best Local Similarity 100.0%; Pred. No. 60;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSVMH 5
 Db 2 SSVMH 6
 |||||

RESULT 6
 S01511
 ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - sea urchin (Strongylocentrotus
 C:Species: Strongylocentrotus purpuratus (purple urchin)
 C:Date: 01-Dec-1989 #sequence_revision 20-Aug-1994 #text_change 03-Jun-2002
 C:Accession: S01511
 R:Jacobs, H.T.; Elliott, D.J.; Math, V.B.; Farquharson, A.
 J. Mol. Biol. 202, 185-217, 1988
 A:Title: Nucleotide sequence and gene organization of sea urchin mitochondrial DNA.
 A:Reference number: S01499; MUID:89011951; PMID:3172215
 A:Accession: S01511
 A:Molecule type: DNA

A:Residues: 1-385 <JAC>
 A:Cross-references: ENBL:X12631
 C:Genetics:

A:Gene: cytb
 A:Genome: mitochondrion
 A:Genetic code: SGC8
 A:Start codon: ATA
 C:Function:

A:Description: the net reaction catalyzed by the ubiquinol-cytochrome-c reductase complex
 1) two hydrogen ions taken up from the mitochondrial matrix and four hydrogen ions relea
 A:Pathway: oxidative phosphorylation; respiratory chain
 C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol-
 C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
 F:17-345/Domain: cytochrome b homology <CBH>
 F:17-215/Domain: cytochrome b6 homology <CB6>
 F:42-58/Domain: transmembrane #status predicted <TM1>
 F:87-105/Domain: transmembrane #status predicted <TM2>
 F:123-139/Domain: transmembrane #status predicted <TM3>
 F:184-206/Domain: transmembrane #status predicted <TM4>
 F:227-345/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
 F:235-251/Domain: transmembrane #status predicted <TM5>
 F:294-310/Domain: transmembrane #status predicted <TM6>
 F:329-349/Domain: transmembrane #status predicted <TM7>
 F:359-375/Domain: transmembrane #status predicted <TM8>
 F:89,188/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
 F:103,202/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 100.0%; Score 25; DB 1; Length 385;
 Best Local Similarity 100.0%; Pred. No. 66;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSVMH 5
 Db 70 SSVMH 74
 |||||

RESULT 7

S04999
 ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - sea lamprey mitochondrion
 C:Species: mitochondrion Petromyzon marinus (sea lamprey)
 C:Date: 06-Oct-1994 #sequence_revision 03-Nov-1995 #text_change 03-Jun-2002
 C:Accession: S54999
 R:Lee, W.J.; Kocher, T.D.
 Genetics 139, 873-887, 1995
 A:Title: Complete sequence of a sea lamprey (Petromyzon marinus) mitochondrial genome: ea
 A:Reference number: S54999; MUID:95229067; PMID:7713438
 A:Accession: S54999
 A:Molecule type: DNA
 A:Residues: 1-396 <LEE>
 A:Cross-references: ENBL:U11880; NID:g515484; PIDN:AAB08737.1; PID:g515485
 A:Note: the authors translated the codon ATA for residue 21 as Ile
 C:Genetics:

A:Gene: cytb
 A:Genome: mitochondrion
 A:Genetic code: SGC1
 C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol-
 C:Keywords: electron transfer; heme; iron; metalloprotein; mitochondrion; oxidative phos
 F:15-343/Domain: cytochrome b homology <CBH>
 F:15-213/Domain: cytochrome b6 homology <CB6>
 F:225-343/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
 F:87,186/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
 F:101,200/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 100.0%; Score 25; DB 2; Length 396;
 Best Local Similarity 100.0%; Pred. No. 68;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSVMH 5
 Db 68 SSVMH 72
 |||||

RESULT 8

```
I37004
Rhesus-like protein - chimpanzee
C:Species: Pan troglodytes (chimpanzee)
C>Date: 04-Oct-1996 #sequence_revision 04-Oct-1996 #text_change 15-Jun-2001
C:Accession: I37004
R:Salvignol, I.; Blancher, A.; Calvas, P.; Clayton, J.; Socha, W.W.; Colin, Y.; Ruffie,
Biochem. Genet. 32, 201-221, 1994
A>Title: Molecular genetics of chimpanzee Rh-related genes: their relationship with the
A:Reference number: I37003; MUID:95085595; PMID:7993375
A:Accession: I37004
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-417 <RES>
A:Cross-references: GB:L37049; NID:g606989; PIDN:AAA65623.1; PID:g606990
C:Superfamily: human erythrocyte membrane protein Rhd

Query Match      100.0%; Score 25; DB 2; Length 417;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
Db 325 SSVMH 329

RESULT 9
I37075
Rhesus-like protein - gorilla
C:Species: Gorilla gorilla (gorilla)
C>Date: 04-Oct-1996 #sequence_revision 04-Oct-1996 #text_change 15-Jun-2001
C:Accession: I37075
R:Salvignol, I.; Blancher, A.; Calvas, P.; Clayton, J.; Socha, W.W.; Colin, Y.; Ruffie,
Biochem. Genet. 32, 201-221, 1994
A>Title: Molecular genetics of chimpanzee Rh-related genes: their relationship with the
A:Reference number: I37003; MUID:95085595; PMID:7993375
A:Accession: I37075
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-417 <RES>
A:Cross-references: GB:L37052; NID:g607009; PIDN:AAA65626.1; PID:g607010
C:Superfamily: human erythrocyte membrane protein Rhd

Query Match      100.0%; Score 25; DB 2; Length 417;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
Db 325 SSVMH 329

RESULT 10
I37076
Rhesus-like protein - gorilla
C:Species: Gorilla gorilla (gorilla)
C>Date: 04-Oct-1996 #sequence_revision 04-Oct-1996 #text_change 15-Jun-2001
C:Accession: I37076
R:Salvignol, I.; Blancher, A.; Calvas, P.; Clayton, J.; Socha, W.W.; Colin, Y.; Ruffie,
Biochem. Genet. 32, 201-221, 1994
A>Title: Molecular genetics of chimpanzee Rh-related genes: their relationship with the
A:Reference number: I37003; MUID:95085595; PMID:7993375
A:Accession: I37076
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-417 <RES>
A:Cross-references: GB:L37053; NID:g607011; PIDN:AAA65627.1; PID:g607012
C:Superfamily: human erythrocyte membrane protein Rhd

Query Match      100.0%; Score 25; DB 2; Length 417;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
```

```
Db 325 SSVMH 329

RESULT 11
I37003
Rhesus-like protein - chimpanzee
C:Species: Pan troglodytes (chimpanzee)
C>Date: 04-Oct-1996 #sequence_revision 04-Oct-1996 #text_change 15-Jun-2001
C:Accession: I37003
R:Salvignol, I.; Blancher, A.; Calvas, P.; Clayton, J.; Socha, W.W.; Colin, Y.; Ruffie,
Biochem. Genet. 32, 201-221, 1994
A>Title: Molecular genetics of chimpanzee Rh-related genes: their relationship with the
A:Reference number: I37003; MUID:95085595; PMID:7993375
A:Accession: I37003
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-417 <RES>
A:Cross-references: GB:L37048; NID:g606987; PIDN:AAA65622.1; PID:g606988
C:Superfamily: human erythrocyte membrane protein Rhd

Query Match      100.0%; Score 25; DB 2; Length 417;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
Db 325 SSVMH 329

RESULT 12
T49849
related to DRPLA protein [imported] - Neurospora crassa
N:Alternate names: protein B24P11.50
C:Species: Neurospora crassa
C>Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
C:Accession: T49849
R:Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, May 2000
A:Reference number: Z25022
A:Accession: T49849
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-538 <SCH>
A:Cross-references: EMBL:AL356833; GSPDB:GN00116; NCSP:B24P11.50
A:Experimental source: BAC clone B24P11; strain OR74A
C:Genetics:
A:Gene: NCSP:B24P11.50
A:Map position: 6
A:Introns: 22/1; 230/2; 492/2

Query Match      100.0%; Score 25; DB 2; Length 538;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
Db 281 SSVMH 285

RESULT 13
A57410
Transcription factor NFATc3 - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 08-Feb-1996 #sequence_revision 08-Feb-1996 #text_change 08-Feb-1996
C:Accession: A57410
R:Ho, S.N.; Thomas, D.J.; Timmerman, L.A.; Li, X.; Francke, U.; Crabtree, G.R.
J. Biol. Chem. 270, 19898-19907, 1995
A>Title: NFATc3, a lymphoid-specific NFATC family member that is calcium-regulated and e
A:Reference number: A57410; MUID:95378239; PMID:7650004
A:Accession: A57410
A>Status: preliminary
A:Molecule type: mRNA
```

A:Residues: 1-1065 <HOA>
A:Cross-references: GB:U28807

Query Match 100.0%; Score 25; DB 2; Length 1065;
Best Local Similarity 100.0%; Pred. No. 2e+02; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0

QY 1 SSVMH 5
|||
Db 773 SSVMH 777

RESULT 14

T42737
gp330 protein precursor - rat
N:Alternate names: megalin
C:Species: Rattus norvegicus (Norway rat)
C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000
C:Accession: T42737
R:Saito, A.; Pietromonaco, S.; Loo, A.K.C.; Farquhar, M.G.
Proc. Natl. Acad. Sci. U.S.A. 91, 9725-9729, 1994
A:Title: Complete cloning and sequencing of rat gp330/megalin, a distinctive member of the gp330 protein family
A:Reference number: A58173; MUID:95024033; PMID:7937880
A:Accession: T42737
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-4660 <SAI>
A:Cross-references: EMBL:L34049; NID:9561852; PID:G561853; PIDN:AAA51369.1
A:Experimental source: strain Sprague-Dawley; kidney
C:Superfamily: alpha-2-macroglobulin receptor; EGF homology; LDL receptor ligand-binding
F:1-25/Domain: signal sequence #status predicted <SIG>
F:26-4660/Product: gp330 protein #status predicted <MAT>

Query Match 100.0%; Score 25; DB 2; Length 4660;
Best Local Similarity 100.0%; Pred. No. 1.1e+03; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0

QY 1 SSVMH 5
|||
Db 954 SSVMH 958

RESULT 15

A82044
DNA-directed RNA polymerase, omega subunit VC2709 [imported] - Vibrio cholerae (strain N16)
C:Species: Vibrio cholerae
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: A82044
R:Heidelberger, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, B.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: A82044
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-90 <HEI>
A:Cross-references: GB:AE004336; GB:AE003852; NID:99657296; PIDN:AAF95849.1; GSPDB:GN00141
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC2709
A:Map position: 1
C:Superfamily: DNA-directed RNA polymerase omega chain

Query Match 96.0%; Score 24; DB 2; Length 90;
Best Local Similarity 80.0%; Pred. No. 24; Indels 0; Gaps 0;
Matches 4; Conservative 1; Mismatches 0

QY 1 SSVMH 5
|||
Db 84 SSVMH 88

RESULT 16

G86154
hypothetical protein T6A9.12 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 17-May-2002
C:Accession: G86154
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Dewar, K.; Ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, R.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: G86154
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-186 <STO>
A:Cross-references: GB:AE005172; NID:99857537; PIDN:AAG00892.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1
C:Superfamily: ADP-ribosylation factor

Query Match 96.0%; Score 24; DB 2; Length 186;
Best Local Similarity 80.0%; Pred. No. 54; Indels 0; Gaps 0;
Matches 4; Conservative 1; Mismatches 0

QY 1 SSVMH 5
|||
Db 31 SSVMH 35

RESULT 17

H86154
hypothetical protein T14P4.20 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 03-Jun-2002
C:Accession: H86154
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Dewar, K.; Ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, R.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: H86154
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-199 <STO>
A:Cross-references: GB:AE005172; NID:99972392; PIDN:AAG10642.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1
C:Superfamily: ADP-ribosylation factor

Query Match 96.0%; Score 24; DB 2; Length 199;
Best Local Similarity 80.0%; Pred. No. 59; Indels 0; Gaps 0;
Matches 4; Conservative 1; Mismatches 0

QY 1 SSVMH 5
|||
Db 31 SSVMH 35

RESULT 18

S57548

hypothetical protein YPR015C - yeast (Saccharomyces cerevisiae)
 N:Alternate names: hypothetical protein LF214C; hypothetical protein YP9531.08C
 C:Species: Saccharomyces cerevisiae
 C:Date: 10-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 19-Apr-2002
 C:Accession: S57548; S59759
 R:Bowman, S.
 submitted to the EMBL Data Library, June 1995
 A:Reference number: S57541
 A:Accession: S57548
 A:Molecule type: DNA
 A:Residues: 1-247 <BOW>
 A:CROSS-references: EMBL:Z49919; NID:G887584; PID:G887592; MIPS:YPR015C
 A:Experimental source: strain A972
 R:Wang, Y.; Ahmed, A.; Bussey, H.; Fortin, N.; Friesen, J.D.; Hall, J.; Storms, R.K.; Vo
 submitted to the EMBL Data Library, July 1995
 A:Description: The sequence of Saccharomyces cerevisiae chromosome XVI right arm.
 A:Reference number: S59746
 A:Accession: S59759
 A:Molecule type: DNA
 A:Residues: 1-247 <WAN>
 A:CROSS-references: EMBL:U31900; NID:G1276597; PID:G939748; MIPS:YPR015C
 C:Genetics:
 A:CROSS-references: SGD:S0006219
 A:Map position: 16R

Query Match 96.0%; Score 24; DB 2; Length 247;
 Best Local Similarity 80.0%; Pred. No. 74;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
 ||:||
 Db 47 SSIMH 51

RESULT 19

T22233
 hypothetical protein F45G2.1 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000
 C:Accession: T22233
 R:Lindsay, S.
 submitted to the EMBL Data Library, March 1997
 A:Reference number: Z19535
 A:Accession: T22233
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-258 <WIL>
 A:CROSS-references: EMBL:Z93382; PIDN:CAB07610.1; GSPDB:GN00021; CESP:F45G2.1
 A:Experimental source: clone F45G2
 C:Genetics:
 A:Gene: CESP:F45G2.1
 A:Map position: 3
 A:Introns: 63/3; 94/1; 134/2; 169/1; 225/3
 C:Superfamily: astacin; astacin homology

Query Match 96.0%; Score 24; DB 2; Length 258;
 Best Local Similarity 80.0%; Pred. No. 78;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
 ||:||
 Db 192 SSIMH 196

RESULT 20

B48826
 high choriolytic hatching proteinase (EC 3.4.24.-) HCE23 precursor - Japanese medaka
 C:Species: Oryzias latipes (Japanese medaka)
 C:Date: 01-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000
 C:Accession: B48826
 R:Yasumasu, S.; Yamada, K.; Akasaka, K.; Mitsunaga, K.; Iuchi, I.; Shimada, H.; Yamagami
 Dev. Biol. 153, 250-258, 1992
 A:Title: Isolation of cDNAs for LCE and HCE, two constituent proteases of the hatching e

A:Reference number: A48826; MUID:93012471; PMID:1397682
 A:Accession: B48826
 A:Molecule type: mRNA; protein
 A:Residues: 1-270 <YAS>
 A:CROSS-references: GB:M96170; NID:G213501; PIDN:AAA49438.1; PID:G213502
 A:Experimental source: orange red variety, embryo
 A:Note: sequence extracted from NCBI backbone (NCBIN:114769, NCBIPI:114770)
 C:Superfamily: astacin; astacin homology
 C:Keywords: glycoprotein; hydrolase; metalloproteinase; zinc
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:21-270/Domain: propeptide #status predicted <PRO>
 F:71-270/Product: low choriolytic hatching proteinase #status predicted <MAT>
 F:88-270/Domain: atactin homology <AST>
 F:53/Binding site: carboxylate (Asn) (covalent) #status predicted
 F:169,173,179,225/Binding site: zinc (His, His, Tyr) #status predicted
 F:170/Active site: Glu #status predicted

Query Match 96.0%; Score 24; DB 2; Length 270;
 Best Local Similarity 80.0%; Pred. No. 82;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
 ||:||
 Db 220 SSIMH 224

RESULT 21

C48826
 high choriolytic hatching proteinase (EC 3.4.24.-) HCE21 precursor - Japanese medaka
 C:Species: Oryzias latipes (Japanese medaka)
 C:Date: 08-Feb-1996 #sequence_revision 08-Feb-1996 #text_change 03-Dec-1999
 C:Accession: C48826
 R:Yasumasu, S.; Yamada, K.; Akasaka, K.; Mitsunaga, K.; Iuchi, I.; Shimada, H.; Yamagami,
 Dev. Biol. 153, 250-258, 1992
 A:Title: Isolation of cDNAs for LCE and HCE, two constituent proteases of the hatching e

A:Reference number: A48826
 A:Accession: C48826
 A:Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-279 <YAS>

A:CROSS-references: GB:M96171
 C:Superfamily: astacin; astacin homology
 C:Keywords: glycoprotein; hydrolase; metalloproteinase; zinc
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:21-279/Domain: propeptide #status predicted <PRO>
 F:80-279/Product: low choriolytic hatching proteinase #status predicted <MAT>
 F:97-279/Domain: atactin homology <AST>
 F:62/Binding site: carboxylate (Asn) (covalent) #status predicted
 F:178,182,188,234/Binding site: zinc (His, His, Tyr) #status predicted
 F:179/Active site: Glu #status predicted

Query Match 96.0%; Score 24; DB 2; Length 279;
 Best Local Similarity 80.0%; Pred. No. 85;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
 ||:||
 Db 229 SSIMH 233

RESULT 22

B69416
 hypothetical protein AFl331 - Archaeoglobus fulgidus
 C:Species: Archaeoglobus fulgidus
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
 C:Accession: B69416
 R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson,
 ; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.;
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
 Nature 380, 364-370, 1997
 A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.N.
 Smith, H.O.; Woesse, C.R.; Venter, J.C.

A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeon
A:Reference number: A69250; MUID:98049343; PMID:9389475
A:Accession: B69416
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-330 <KLE>
A:Cross-references: GB:AE001012; GB:AE000782; NID:g2689335; PIDN:AAB99926.1; PID:g264925

Query Match 96.0%; Score 24; DB 2; Length 330;
Best Local Similarity 80.0%; Pred. No. 1e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
DB 155 SSIH 159

RESULT 23
C87194
probable transmembrane transport protein ML2279 [imported] - Mycobacterium leprae
C:Species: Mycobacterium leprae
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C:Accession: C87194
R: Cole, S.T.; Eiglmeyer, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Holt,
R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd,
eam, M.A.; Rutherford, K.M.
Nature 409, 1007-1011, 2001
A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Seeger,
A:Title: Massive gene decay in the leprosy bacillus.
A:Reference number: A86909; MUID:21128732; PMID:11234002
A:Accession: C87194
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-496 <STO>
A:Cross-references: GB:AL450380; NID:gl3093915; PIDN:CAC31795.1; GSPDB:GN00147
C:Genetics:
A:Gene: ML2279

Query Match 96.0%; Score 24; DB 2; Length 496;
Best Local Similarity 80.0%; Pred. No. 1.6e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
DB 305 SSIH 309

RESULT 24
AD3409
ABC transporter ATP-binding protein BMEI1258 [imported] - Brucella melitensis (strain 16
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 15-Feb-2002
C:Accession: AD3409
R: DelVecchio, V.G.; Kaputral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
Mazur, M.; Goltzman, E.; Selkov, E.; Elizer, P.H.; Hagius, S.; O'Callaghan, D.; Letes
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A:Reference number: AD3252; PMID:11756688
A:Accession: AD3409
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-646 <KUR>
A:Cross-references: GB:AE008917; PIDN:AAL52439.1; PID:gl7983243; GSPDB:GN00190
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BMEI1258
A:Map position: I
C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

Query Match 96.0%; Score 24; DB 2; Length 646;
Best Local Similarity 80.0%; Pred. No. 2.2e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
DB 224 SSIH 228

RESULT 25

T30018
hypothetical protein F38E9.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T30018
R: Wu, X.; Gattung, S.
submitted to the EMBL Data Library, January 1996
A:Description: The sequence of C. elegans cosmid F38E9.
A:Reference number: Z20722
A:Accession: T30018
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-767 <WUX>
A:Cross-references: EMBL:U46668; PIDN:AAA93348.1; CESP:F38E9.2
C:Genetics:
A:Gene: CESP:F38E9.2
A:Introns: 13/1; 37/3; 62/1; 89/1; 127/1; 164/2; 182/1; 235/3; 269/3; 321/2; 360/1; 455/3

Query Match 96.0%; Score 24; DB 2; Length 767;
Best Local Similarity 80.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
DB 36 SSIH 40

Search completed: September 24, 2004, 01:54:33
Job time : 12.4528 secs

This page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2005, 17:57:45 ; Search time 37.6562 Seconds
(without alignments)
51.354 Million cell updates/sec

Title: US-10-088-639a-2_COPY_158_162

Perfect score: 25

Sequence: 1 SSMVH 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

1: Geneseqp1980s:*

2: Geneseqp1990s:*

3: Geneseqp2000s:*

4: Geneseqp2001s:*

5: Geneseqp2002s:*

6: Geneseqp2003as:*

7: Geneseqp2003bs:*

8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	25	100.0	12	8	ADL73176 CDR1 of h
2	25	100.0	27	8	ADQ26158 INSP005B
3	25	100.0	27	8	ADQ26138 INSP005A
4	25	100.0	69	5	ABP33613 Human ORP
5	25	100.0	76	4	AAM57202 Human bra
6	25	100.0	76	4	ABG51282 Human liv
7	25	100.0	96	4	ABBI1748 Human odo
8	25	100.0	112	4	AAE12026 Human G-p
9	25	100.0	117	4	AACT2494 Human OR-
10	25	100.0	117	4	AACT2494 Human OR-
11	25	100.0	133	4	ABG10765 Novel hum
12	25	100.0	137	8	ADL73173 Anti-idio
13	25	100.0	144	7	ADC97464 E. faeciu
14	25	100.0	167	4	ABR343245 Amino aci
15	25	100.0	248	6	ABR343245 Human PMW
16	25	100.0	249	4	AAE68087 An anti-a
17	25	100.0	253	5	ABG66935 Novel G-p
18	25	100.0	259	7	ADB47696 A. gossyp
19	25	100.0	259	8	ADH42275 Novel hum
20	25	100.0	259	8	ADH42277 Novel hum
21	25	100.0	259	8	ADH42271 Novel hum
22	25	100.0	259	8	ADH42273 Novel hum
23	25	100.0	282	7	AAO23391 Human 656
24	25	100.0	286	7	ADB47698 A. gossyp
25	25	100.0	300	5	ABB07687 Human pro

26	25	100.0	300	6	AAE35322 Human sec
27	25	100.0	300	8	ADQ26167 INSP005 p
28	25	100.0	300	8	ADQ26171 INSP005 f
29	25	100.0	306	4	AAE11905 Human G-p
30	25	100.0	313	4	AAE11905 Human G-p
31	25	100.0	314	4	AAE11906 Human G-p
32	25	100.0	314	4	AAE11906 Human G-p
33	25	100.0	314	4	AAE11906 Human G-p
34	25	100.0	314	4	AAE11906 Human G-p
35	25	100.0	314	4	AAE11906 Human G-p
36	25	100.0	314	4	AAE11906 Human G-p
37	25	100.0	314	5	AAE11906 Human G-p
38	25	100.0	314	5	AAE11906 Human G-p
39	25	100.0	314	5	AAE11906 Human G-p
40	25	100.0	314	5	AAE11906 Human G-p
41	25	100.0	314	5	AAE11906 Human G-p
42	25	100.0	314	5	AAE11906 Human G-p
43	25	100.0	314	5	AAE11906 Human G-p
44	25	100.0	314	7	ADC86303 Human GPC
45	25	100.0	314	7	ADC86305 Human GPC

ALIGNMENTS

RESULT 1

ADL73176
ID ADL73176 standard; peptide; 12 AA.

XX AC ADL73176;

XX DT 20-MAY-2004 (first entry)

XX DE CDR1 of heavy chain of anti-idiotypic antibody 14C12.

XX KW mouse; anti-idiotypic antibody; 14C12; Factor VIII inhibitory antibody;
XX KW C2 domain; Factor VIII; procoagulant; bleeding disorder; haemophilia;
XX KW B cell apoptosis; antibody; heavy chain; CDR.
XX OS Mus musculus.
XX PN EPI388544-A1.
XX PD 11-FEB-2004.
XX PF 31-JUL-2002; 2002EP-00447150.
XX PR 31-JUL-2002; 2002EP-00447150.
(COLL-) COLLEN RES FOUND VZW D.
Gille JGG, Saint-Remy JR, Jacquemin MG;
WPI; 2004-158719/16.
New monoclonal anti-idiotypic antibodies against human Factor VIII
inhibitory antibodies, useful for treating or preventing bleeding
disorders of hemophilia patient.
Claim 8; Page 15; 25pp; English.
ADL73176-ADL73181 represent complementarity determining regions (CDRs)
derived from the heavy (ADL73176-ADL73178) and the light (ADL73179-
ADL73181) chains of murine anti-idiotypic antibody 14C12. This antibody
is directed against a human Factor VIII inhibitory antibody directed
towards the C2 domain of Factor VIII. The anti-idiotypic antibody is
characterized by having the capacity to neutralize by at least 50% the
inhibition of FVIII procoagulant activity mediated by inhibitory
antibodies against the C2 domain of FVIII. The anti-idiotypic antibodies
of the invention may be used to treat, prevent or reduce bleeding
disorders of haemophilia patients, or to induce apoptosis of B cells
carrying anti-C2 inhibitory antibodies in a haemophilia patient.

```

SQ      Sequence 12 AA;
Query Match      100.0%; Score 25; DB 8; Length 12;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SSVMH 5
        |||||
DB      6 SSVMH 10

RESULT 2
ADQ26158
ID      ADQ26158 standard; peptide; 27 AA.
XX
AC      ADQ26158;
XX
DT      23-SEP-2004 (first entry)
XX
DE      INSP005B protein sequence exon 7.
XX
KW      secreted; INSP005A; INSP005B; metalloprotease;
KW      choriolysin/astacin-like metalloprotease; respiratory disorder;
KW      emphysema; cystic fibrosis; metabolic disorder; cardiovascular disorder;
KW      bacterial infection; hypertension; proliferative disorder; cancer;
KW      autoimmune; inflammatory disorder; rheumatoid arthritis;
KW      neurological disorder; developmental disorder; reproductive disorder.
XX
OS      Homo sapiens.
XX
FN      WO2004056983-A2.
XX
PD      08-JUL-2004.
XX
PF      23-DEC-2003; 2003WO-GB005664.
XX
PR      23-DEC-2002; 2002GB-00030006.
XX
PA      (ARES-) ARES TRADING SA.
XX
PI      Fagan RJ, Phelps CB, Power C, Mitter RJ, Boschert U, Chvatchko Y;
DR      WPI: 2004-507715/48.
DR      N-PSDB; ADQ26157.
XX
XX      New metalloprotease proteins (INSP005a or INSP005b) for diagnosing,
PT      preventing or treating diseases associated with metalloproteases, e.g.
PT      emphysema, cystic fibrosis, metabolic disorders, cardiovascular disorders
PT      or cancer.
XX
PS      Example 2; SEQ ID NO 28; 101pp; English.
XX
CC      This sequence represents a fragment of a new secreted polypeptide,
CC      INSP005B. This protein is a metalloprotease, specifically a
CC      choriolysin/astacin-like metalloprotease. The INSP005 polypeptides and
CC      nucleic acid molecules of the invention are useful in the therapy or
CC      diagnosis of diseases or in the manufacture of a medicament for the
CC      treatment of a disease, such as a respiratory disorder, including
CC      emphysema and cystic fibrosis, a metabolic disorder, a cardiovascular
CC      disorder, a bacterial infection, hypertension, a proliferative disorder,
CC      including cancer, an autoimmune/inflammatory disorder, including
CC      rheumatoid arthritis, a neurological disorder, a developmental disorder,
CC      a reproductive disorder or other pathological condition in which
CC      metalloproteases are implicated.
XX
SQ      Sequence 27 AA;
Query Match      100.0%; Score 25; DB 8; Length 27;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SSVMH 5
        |||||
DB      6 SSVMH 10

RESULT 3
ADQ26138
ID      ADQ26138 standard; peptide; 27 AA.
XX
AC      ADQ26138;
XX
DT      23-SEP-2004 (first entry)
XX
DE      INSP005A protein sequence exon 4.
XX
KW      secreted; INSP005A; INSP005B; metalloprotease;
KW      choriolysin/astacin-like metalloprotease; respiratory disorder;
KW      emphysema; cystic fibrosis; metabolic disorder; cardiovascular disorder;
KW      bacterial infection; hypertension; proliferative disorder; cancer;
KW      autoimmune; inflammatory disorder; rheumatoid arthritis;
KW      neurological disorder; developmental disorder; reproductive disorder.
XX
OS      Homo sapiens.
XX
FN      WO2004056983-A2.
XX
PD      08-JUL-2004.
XX
PF      23-DEC-2003; 2003WO-GB005664.
XX
PR      23-DEC-2002; 2002GB-00030006.
XX
PA      (ARES-) ARES TRADING SA.
XX
PI      Fagan RJ, Phelps CB, Power C, Mitter RJ, Boschert U, Chvatchko Y;
DR      WPI: 2004-507715/48.
DR      N-PSDB; ADQ26137.
XX
XX      New metalloprotease proteins (INSP005a or INSP005b) for diagnosing,
PT      preventing or treating diseases associated with metalloproteases, e.g.
PT      emphysema, cystic fibrosis, metabolic disorders, cardiovascular disorders
PT      or cancer.
XX
PS      Example 2; SEQ ID NO 8; 101pp; English.
XX
CC      This sequence represents a fragment of a new secreted polypeptide,
CC      INSP005A. This protein is a metalloprotease, specifically a
CC      choriolysin/astacin-like metalloprotease. The INSP005 polypeptides and
CC      nucleic acid molecules of the invention are useful in the therapy or
CC      diagnosis of diseases or in the manufacture of a medicament for the
CC      treatment of a disease, such as a respiratory disorder, including
CC      emphysema and cystic fibrosis, a metabolic disorder, a cardiovascular
CC      disorder, a bacterial infection, hypertension, a proliferative disorder,
CC      including cancer, an autoimmune/inflammatory disorder, including
CC      rheumatoid arthritis, a neurological disorder, a developmental disorder,
CC      a reproductive disorder or other pathological condition in which
CC      metalloproteases are implicated.
XX
SQ      Sequence 27 AA;
Query Match      100.0%; Score 25; DB 8; Length 27;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SSVMH 5
        |||||
DB      20 SSVMH 24

RESULT 4
ABP33613
ID      ABP33613 standard; protein; 69 AA.
XX
AC      ABP33613;

```

XX DT 08-JUL-2002 (first entry)
XX DE Human ORF2586 protein, SEQ ID NO:5172.
XX KW Human; ORF; open reading frame; ORFX; drug screening; diagnosis;
KW disease monitoring; cytokine; cell proliferation; cell differentiation;
KW immune modulation; haematopoiesis regulation; tissue growth;
KW angiogenesis; activin; inhibin; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; tumour inhibition; bodily characteristics; fertility;
KW behaviour; cancer; proliferative disorder; neurological disorder;
KW cardiovascular disease; immune system disorder; organ transplantation;
KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
KW hypothyroidism; cholesterol ester storage disease; infection; vulnary;
KW vasotropic; antipsoriatic; antidiabetic; cytostatic; nootropic;
KW neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;
KW cardiant; hypotensive; antihypertoid; antiinflammatory; immunomodulator;
KW dermatological; analgesic; virucide; antibacterial; fungicide.
XX OS Homo sapiens.
XX AC WO200190366-A2.
XX PN 29-NOV-2001.
XX DT 24-MAY-2001; 2001WO-US017076.
XX PF 24-MAY-2000; 2000US-0206690P.
XX PR (CURA-) CURAGEN CORP.
XX PA Leach MD, Shinkets RA;
XX PI WPI: 2002-106200/14.
XX DR N-PSDB; ABN77639.
XX DT Novel human polypeptides and polynucleotides useful for diagnosing,
PT preventing and treating cardiovascular disease, neurodegenerative,
PT hyperproliferative disorders and disorders related to organ
PT transplantation.
XX Claim 10; Page 1545; 2508pp; English.
XX Sequences ABP31028-ABP35561 represent 4534 novel human proteins
CC designated ORF (open reading frame) 1-4534, and sequences ABN75054-
CC ABN79587 represent cDNAs encoding them. The invention also encompasses
CC polypeptides at least 80% identical to the ORF1-ORF4534 collectively
CC referred to as ORFX) proteins, polynucleotides at least 85% identical to
CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
CC polynucleotides, the recombinant production of ORFX proteins, antibodies
CC specific for ORFX proteins, methods of detecting ORFX polynucleotides and
CC polypeptides, methods of screening for modulators of ORFX expression or
CC activity, and methods of screening individuals for a predisposition to an
CC ORFX-associated disorder. The ORFX proteins of the invention have a wide
CC range of biological activities, such as cytokines, cell proliferation,
CC cell differentiation, immune modulation, haematopoiesis regulation,
CC tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
CC receptor/ligand, antiinflammatory activity, tumour inhibition activity,
CC and antineoplastic activity, and may also be involved in the determination
CC of bodily characteristics, fertility and behaviour. ORFX proteins,
CC nucleic acids and antibodies may be used in the treatment of cancers,
CC other proliferative disorders such as psoriasis and benign tumours,
CC neurological disorders such as epilepsy and Alzheimer's disease,
CC cardiovascular diseases, immune system disorders, disorders related to
CC organ transplantation, disorders of tissue growth and regeneration,
CC diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester
CC storage disease, and infectious diseases caused by viral, bacterial,
CC fungal and other pathogens. ORFX nucleic acids may also be used as a
CC source of primers and probes, in the detection of ORFX genomic sequences
CC or transcripts, in the identification and cloning of homologous
CC sequences, in genetic diagnosis, and in forensic biology. The ORFX
CC nucleic acids may additionally be used to produce transgenic animals

CC which may be useful for studying the function and/or activity of ORFX
CC protein, and in drug screening. The ORFX proteins may also be used as
CC immunogens to generate specific antibodies, which are useful in the
CC diagnosis, treatment and monitoring of ORFX-associated diseases
XX SQ Sequence 69 AA;
Query Match 100.0%; Score 25; DB 5; Length 69;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 SSVMH 5
Db 23 SSVMH 27
RESULT 5
AAM57202
ID AAM57202 standard; protein; 76 AA.
XX AC AAM57202;
XX DT 05-NOV-2001 (first entry)
XX DE Human brain expressed single exon probe encoded protein SEQ ID NO: 29307.
XX KW Human; brain expressed exon; gene expression analysis; probe; microarray;
KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.
XX OS Homo sapiens.
XX PN WO200157275-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000667.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-483446/52.
XX Single exon nucleic acid probes for analyzing gene expression in human
PT brains.
XX Example 4; SEQ ID NO 29307; 650pp + Sequence Listing; English.
XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is a protein encoded by one of
CC the probes of the invention
XX SQ Sequence 76 AA;
Query Match 100.0%; Score 25; DB 4; Length 76;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 SSVMH 5
Db 35 SSVMH 39

RESULT 6
 ABG51282
 ID ABG51282 standard; peptide; 76 AA.
 AC ABG51282;
 XX
 DT 25-FEB-2003 (first entry)
 XX
 DE Human liver peptide, SEQ ID NO 29930.
 XX
 KW Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
 KW hypercholesterolaemia; coronary heart disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200157273-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US000664.
 XX
 PR 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-488998/53.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for analyzing
 PT gene expression in human adult liver.
 XX
 PS Claim 27; SEQ ID NO 29930; 658pp; English.
 XX
 CC The invention relates to a single exon nucleic acid probe (SENP) (I) for
 CC measuring human gene expression in a sample derived from human adult
 CC liver, comprising one of 13109 defined nucleotide sequences given in the
 CC specification (or complements/ fragments). The probe hybridises at high
 CC stringency to a nucleic acid molecule expressed in the human adult liver.
 CC (I) may be used for predicting, measuring and displaying gene expression
 CC in samples derived from human adult liver. The genes identified may be
 CC involved in genetic liver diseases such as cirrhosis,
 CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
 CC associated with coronary heart disease. ABG47348-ABG59930 represent human
 CC liver single exon encoded peptides of the invention. Note: The sequence
 CC information for this patent does not appear in the printed specification
 CC but was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 76 AA;
 Query Match 100.0%; Score 25; DB 4; Length 76;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02; Indels 0; Gaps 0;
 Matches 5; Conservative 0; Mismatches 0;
 QY 1 SSVMH 5
 |||||
 Db 35 SSVMH 39
 RESULT 7
 ABB11748
 ID ABB11748 standard; peptide; 96 AA.
 XX
 AC ABB11748;

XX
 DT 11-JAN-2002 (first entry)
 XX
 DE Human odorant receptor S18 homologue, SEQ ID NO:2118.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
 KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
 KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
 KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
 KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
 KW chronic inflammatory condition; proliferative retinopathy;
 KW atherosclerosis; coronary heart disease; arterial ischaemia;
 KW bone disorder; osteoporosis; vascular growth disorder;
 KW tissue regeneration; wound healing; infection; immune disorder;
 KW cell culture; drug screening; gene therapy; antiinflammatory;
 KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
 KW cytosstatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
 KW antifungal; vulnery; antiulcer.
 XX
 OS Homo sapiens.
 XX
 PN WO200157188-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 05-FEB-2001; 2001WO-US0003800.
 XX
 PR 03-FEB-2000; 2000US-00496914.
 PR 27-APR-2000; 2000US-00560875.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;
 XX
 DR WPI; 2001-457740/49.
 XX
 PT Human proteins and DNA encoding sequences useful for preventing, treating
 PT or ameliorating a medical condition in a mammalian subject e.g. arthritis
 PT and cancer.
 XX
 PS Claim 20; Page 238; 1963pp; English.
 XX
 CC Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
 CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
 CC invention also relates to vectors and recombinant host cells comprising a
 CC nucleotide of the invention, methods of producing the novel polypeptides,
 CC antibodies against the polypeptides, methods of detecting the nucleotides
 CC or polypeptides in a sample, and methods of identifying compounds which
 CC bind to polypeptides of the invention. Although novel, many of the
 CC polypeptides of the invention have homology to known proteins, thereby
 CC giving an insight into their probable biological activities, and hence
 CC potential therapeutic applications. The polypeptides of the invention may
 CC have various activities, including cytokine, cell proliferation or cell
 CC differentiation activities; stem cell growth factor activity;
 CC haematopoiesis regulatory activity; tissue growth activity;
 CC immunomodulatory activity; activin- or inhibin-related activities;
 CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
 CC thrombolytic activities; receptor or ligand activities; or may be
 CC involved in oncogenesis, cancer cell proliferation or metastasis.
 CC Depending on their biological activities, polypeptides and nucleotides of
 CC the invention are useful for preventing, treating or ameliorating medical
 CC conditions, e.g., by protein or gene therapy. Such conditions include
 CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
 CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
 CC proliferative retinopathy, atherosclerosis, coronary heart disease,
 CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
 CC vascular growth. Polypeptides involved with tissue regeneration and
 CC repair (or nucleic acids encoding them) may be used to promote wound
 CC healing (e.g., of burns, incisions and ulcers), while those with
 CC immunomodulatory activities may be used in the treatment of viral,
 CC bacterial and fungal infections in addition to immune disorders.
 CC Polypeptides with growth factor activity may be used in cell cultures to

CC promote cell growth. For example, such polypeptides may be used to
 CC manipulate stem cells in culture to give rise to neuroepithelial cells
 CC that can be used to augment or replace cells damaged by illness.
 CC autoimmune disease or accidental damage. The polypeptides and nucleotides
 CC may also be used in the diagnosis of the above conditions, and in drug
 CC screening techniques. The present sequence represents a novel human
 CC polypeptide of the invention
 XX
 SQ Sequence 96 AA;
 Query Match 100.0%; Score 25; DB 4; Length 96;
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SSVHM 5
 DB 39 SSVHM 43
 |||||
 RESULT 8
 AAE12026
 ID AAE12026 standard; protein; 112 AA.
 XX
 AC AAE12026;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Human G-protein coupled receptor, GCRC-5.
 XX
 KW Human; G-protein coupled receptor; GCRC; gene therapy;
 KW proteosome analysis; cell proliferative disorder; arteriosclerosis;
 KW cancer; neurological disorder; Huntington's disease; Parkinson's disease;
 KW cardiovascular disorder; atherosclerosis; congestive heart failure;
 KW gastrointestinal disorder; gastritis; nausea; autoimmune; anaemia;
 KW inflammatory disorder; acquired immunodeficiency syndrome; AIDS;
 KW metabolic disorder; diabetes; obesity; viral infection; drug screening;
 KW chromosomal mapping; cytostatic; immunomodulatory; antiinflammatory;
 KW viricide.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..27 /label= Signal_peptide
 FT Domain 1..23 /label= Transmembrane_domain
 FT Protein 28..112
 FT Binding-site 64..112 /note= "Mature human GCRC-5 protein"
 FT Domain 81..97 /note= "Visual pigments retinal binding site"
 FT /label= G-protein_coupled_receptor_domain
 XX
 PN WO200172836-A2.
 XX
 PD 04-OCT-2001.
 XX
 PF 29-MAR-2001; 2001WO-US010436.
 XX
 PR 29-MAR-2000; 2000US-0193051P.
 PR 06-APR-2000; 2000US-0195155P.
 PR 20-APR-2000; 2000US-0199084P.
 PR 28-APR-2000; 2000US-0200551P.
 PR 05-MAY-2000; 2000US-020278P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Policky JL, Tribouley CM, Tang YT, Baughn MR, Graul R, Khan EA;
 PI Nguyen DB, Patterson C, Lal P, Au-Young J, Yang J, Hafalia A;
 PI Wallia NK, Das D;
 XX
 DR WPI; 2001-616472/71.
 DR N-PSDB; AAD19581.

XX New human G-protein coupled receptors, useful for treatment and diagnosis
 PT of e.g. cell proliferation, also screening for specific modulators, and
 PT related nucleic acid.
 XX
 PS Claim 1; Page 106; 111pp; English.
 XX
 CC The present sequence is a human G-protein coupled receptor, GCRC-5
 CC protein. The GCRCs are used for treating or preventing disorders
 CC associated with decreased expression of functional GCRC, and for
 CC identifying specific agonists and antagonists, also binding agents and
 CC modulators. They can also be used for generating specific antibodies and
 CC for proteosome analysis. Disorders that can be treated include cell
 CC proliferative disorders, e.g., arteriosclerosis and cancer, neurological
 CC disorders, e.g., Huntington's disease and Parkinson's disease,
 CC cardiovascular disorders, e.g., atherosclerosis and congestive heart
 CC failure, gastrointestinal disorders, e.g., gastritis and nausea,
 CC autoimmune/inflammatory disorders, e.g., acquired immunodeficiency
 CC syndrome (AIDS) and anaemia, metabolic disorders, e.g., diabetes and
 CC obesity and viral infections. Nucleic acids that encode GCRC are used
 CC for identifying agents that alter its expression, for assessing toxicity
 CC of test compounds, and as sources of primers and probes for diagnostic
 CC detection of GCRC DNA and of therapeutic antisense and ribozyme
 CC sequences. They can also be used in gene therapy, for chromosomal
 CC mapping, and for recombinant production of GCRC. The antibodies are
 CC useful for diagnosis and monitoring of diseases associated with GCRC
 CC expression, for detecting and purifying GCRC, and as therapeutic agents
 CC and for drug screening
 XX
 SQ Sequence 112 AA;
 Query Match 100.0%; Score 25; DB 4; Length 112;
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SSVHM 5
 DB 55 SSVHM 59
 |||||
 RESULT 9
 AAG72494
 ID AAG72494 standard; protein; 117 AA.
 XX
 AC AAG72494;
 XX
 DT 30-JUL-2001 (first entry)
 XX
 DE Human OR-like polypeptide query sequence, SEQ ID NO: 2175.
 XX
 KW Human; olfactory receptor; OR; primary scent determination;
 KW secondary scent determination; polypeptide library; odour receptor;
 KW scent profile; scent fingerprint; scent representation.
 XX
 OS Homo sapiens.
 XX
 PN WO200127158-A2.
 XX
 PD 19-APR-2001.
 XX
 PF 06-OCT-2000; 2000WO-US027582.
 XX
 PR 08-OCT-1999; 99US-0158615P.
 PR 24-FEB-2000; 2000US-0184803P.
 XX
 PA (DIGI-) DIGISCENTS.
 PA (YEDA) YEDA RES & DEV CO LTD.
 XX
 PI Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;
 DR WPI; 2001-290713/30.
 XX
 PT New polynucleotides which encode polypeptides involved in olfactory

PT sensation for identifying olfactory agonists and antagonists.
 PS Example 6; Page 1464-1465; 1857pp; English.
 XX

CC The present sequence is a polypeptide encoded by one of 344 newly mined
 CC human genes. It was used as a query sequence in a database search of
 CC olfactory receptor (OR)-like sequences. The invention relates to isolated
 CC polynucleotides encoding polypeptides involved in olfactory sensation.
 CC The polynucleotides can be used in screening for olfactory agonists and
 CC antagonists. The methods allow for the determination of primary scents
 CC and the identification of the odour receptors used to detect these
 CC primary scents. The methods also enable determination of secondary scents
 CC and the identification of combinations of odour receptors that are
 CC involved in detecting such secondary scents. This enables the
 CC construction of a scent representation (also called a scent fingerprint
 CC or scent profile), which may be used to re-create and edit scents.
 CC Libraries of olfactory receptors are useful for determining the
 CC interaction pattern of a composition with the receptors, and can be used
 CC for determining differences in the olfactory faculties of different
 CC individuals
 XX

SQ Sequence 117 AA;
 Query Match 100.0%; Score 25; DB 4; Length 117;
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSVMH 5
 |||||
 Db 63 SSVMH 67

RESULT 10
 AAG71795
 ID AAG71795 standard; protein; 117 AA.
 AC AAG71795;
 XX

DT 30-JUL-2001 (first entry)
 DE Human olfactory receptor polypeptide, SEQ ID NO: 1476.
 KW Human; olfactory receptor; OR; primary scent determination;
 KW secondary scent determination; polypeptide library; odour receptor;
 KW scent profile; scent fingerprint; scent representation.
 XX

OS Homo sapiens.
 PN WO200127158-A2.
 XX

PD 19-APR-2001.
 XX

PF 06-OCT-2000; 2000WO-US027582.
 XX

PR 08-OCT-1999; 99US-0158615P.
 PR 24-FEB-2000; 2000US-0184809P.
 XX

PA (DIGI-) DIGISCENTS.
 PA (YEDA) YEDA RES & DEV CO LTD.
 XX

PI Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;
 XX

DR WPI; 2001-290713/30.
 XX

PT New polynucleotides which encode polypeptides involved in olfactory
 PT sensation for identifying olfactory agonists and antagonists.
 XX

PS Claim 11; Page 934-935; 1857pp; English.
 XX

CC The present sequence is an olfactory receptor which is encoded by one of
 CC a number of novel polynucleotides. The polynucleotides can be used in
 CC screening for olfactory agonists and antagonists. The methods allow for
 CC the determination of primary scents and the identification of the odour
 CC

CC receptors used to detect these primary scents. The methods also enable
 CC determination of secondary scents and the identifying such secondary scents.
 CC This enables the construction of a scent representation (also called a
 CC scent fingerprint or scent profile), which may be used to re-create and
 CC edit scents. Libraries of olfactory receptors are useful for determining
 CC the interaction pattern of a composition with the receptors, and can be
 CC used for determining differences in the olfactory faculties of different
 CC individuals
 XX

SQ Sequence 117 AA;
 Query Match 100.0%; Score 25; DB 4; Length 117;
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSVMH 5
 |||||
 Db 63 SSVMH 67

RESULT 11
 ABG10765
 ID ABG10765 standard; protein; 133 AA.
 AC ABG10765;
 XX

DT 13-FEB-2002 (first entry)
 DE Novel human diagnostic protein #10756.
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX

OS Homo sapiens.
 PN WO200175067-A2.
 XX

PD 11-OCT-2001.
 XX

PF 30-MAR-2001; 2001WO-US008631.
 XX

PR 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 XX

PA (HYSE-) HYSEQ INC.
 XX

PI Drmanac RT, Liu C, Tang YT;
 XX

DR WPI; 2001-639362/73.
 DR N-PSDB; AAS74952.
 XX

XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX

PS Claim 20; SEQ ID NO 41124; 103pp; English.
 XX

CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC

CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 133 AA;

Query Match 100.0%; Score 25; DB 4; Length 133;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSVMH 5
|
|
|
|
|
DB 12 SSVMH 16

RESULT 12

ADL73173
ID ADL73173 standard; protein; 137 AA.

XX
AC ADL73173;

XX 20-MAY-2004 (first entry)

XX Anti-idiotypic antibody 14C12 heavy chain variable region.

XX Heavy chain; variable region; mouse; anti-idiotypic antibody; 14C12;
KW Factor VIII inhibitory antibody; C2 domain; Factor VIII; procoagulant;
KW bleeding disorder; haemophilia; B cell apoptosis; antibody.

XX OS Mus musculus.

XX Key Location/Qualifiers

FT Region 26..37
/note= "complementarity determining region 1"
FT Region 49..66
/note= "complementarity determining region 2"
FT Region 99..111
/note= "complementarity determining region 3"

XX EP1388544-A1.

XX 11-FEB-2004.

XX 31-JUL-2002; 2002EP-00447150.

XX 31-JUL-2002; 2002EP-00447150.

XX (COLL-) COLLEN RES FOUND VZW D.

XX Gilie JGG, Saint-Remy JR, Jacquemin MG;

XX WPI; 2004-158719/16.

XX N-PSDB; ADL73172.

XX New monoclonal anti-idiotypic antibodies against human Factor VIII
PT inhibitory antibodies, useful for treating or preventing bleeding
PT disorders of hemophilia patient.

XX Disclosure; Fig 3A; 25pp; English.

XX The present sequence represents the heavy chain variable region of murine
CC anti-idiotypic antibody 14C12. This antibody is directed against a human
CC Factor VIII inhibitory antibody directed towards the C2 domain of Factor
CC VIII. The anti-idiotypic antibody is characterized by comprising
CC complementarity determining regions (CDRs) given in ADL73176-ADL73181,
CC and by having the capacity to neutralize by at least 50% the inhibition
CC of FVIII procoagulant activity mediated by inhibitory antibodies against
CC the C2 domain of FVIII. The anti-idiotypic antibodies of the invention
CC may be used to treat, prevent or reduce bleeding disorders of haemophilia

CC patients, or to induce apoptosis of B cells carrying anti-C2 inhibitory
CC antibodies in a haemophilia patient.

XX Sequence 137 AA;

Query Match 100.0%; Score 25; DB 8; Length 137;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSVMH 5
|
|
|
|
|
DB 31 SSVMH 35

RESULT 13

ADC97464
ID ADC97464 standard; protein; 144 AA.

XX
AC ADC97464;

XX 01-JAN-2004 (first entry)

XX E. faecium protein sequence SEQ ID 7091.

XX Vaccine; urinary tract infection; bacteraemia; endocarditis; wound;
KW abdominal-pelvic infection.

XX Enterococcus faecium.

XX US583275-B1.

XX 24-JUN-2003.

XX 30-JUN-1998; 98US-00107532.

XX 02-JUL-1997; 97US-0051571P.

XX 14-MAY-1998; 98US-0085598P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Doucette-Stamm LA, Bush D;

XX WPI; 2003-799836/75.

XX N-PSDB; ADC93810.

XX New isolated nucleic acid derived from Enterococcus faecium encoding an
PT Enterococcus faecium polypeptide useful for detection, prevention and
PT treatment of a pathological condition resulting from a bacterial
PT infection.

XX Example 1; SEQ ID NO 7091; 243pp; English.

XX The invention relates to an isolated nucleic acid derived from
CC Enterococcus faecium encoding an Enterococcus faecium polypeptide having
CC one of 10 fully defined sequences given in the (or comprising 40
CC sequential nucleotides chosen from any of the nucleic acids, its
CC complement or sequences hybridizing to it). Also included are a
CC recombinant vector comprising the nucleic acid operably linked to
CC transcription regulatory element, a cell comprising the vector and a
CC single-stranded probe comprising the nucleic acid. The nucleic acids are
CC chosen from 3654 disclosed sequences encoding 3654 disclosed proteins.

XX The nucleic acids are useful for diagnosing pathological conditions
CC resulting from E. faecium bacterial infection (e.g. urinary tract
CC infection, bacteraemia, endocarditis, wounds and abdominal-pelvic
CC infection) and for screening drugs such as agonists and antagonists. The
CC nucleic acid is useful for recombinant production of Candida albicans -
CC derived peptides or antisense polypeptides. Pharmaceutical compositions
CC and vaccines containing the nucleic acid are useful for preventing or
CC treating Enterococcus faecium infections. The present sequence represents
XX one if the disclosed E. faecium proteins.

XX Sequence 144 AA;

Query Match 100.0%; Score 25; DB 7; Length 144;
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
 DB 94 SSVMH 98
 |||||

RESULT 14

AA85005
 ID AAB85005 standard; protein; 167 AA.

XX
 AC AAB85005;

DT 06-AUG-2001 (first entry)

DE Amino acid sequence of A1138218, a PHOR-1 family member.

XX G-protein-coupled receptor; prostate; cancer; PHOR-1; kidney; uterine;
 KW cervical; stomach; rectal; cystostatic; vaccine; cell function regulator;
 KW human; prostate homologue of olfactory receptor-1.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 1..167

FT /note= "residues Xaa are encoded by internal stop codons"

PN WO200125434-A1.

XX 12-APR-2001.

XX 05-OCT-2000; 2000WO-US027543.

XX 05-OCT-1999; 99US-0157902P.

XX (UROC-) UROGENESYS INC.

XX Raitano AB, Afar DEH, Jakobovits A, Faris M, Hubert RS;

XX Mitchell SC, Saffran DC;

XX WPI; 2001-367230/38.

XX N-PSDB; AAF83882.

XX Novel gene designated PHOR-1, a G-protein-coupled receptor up-regulated
 PT in prostate cancer, useful as diagnostic marker and therapeutic target
 PT for cancers of prostate, kidney, uterus.

XX Disclosure; Fig 22; 139pp; English.

XX The invention relates to a novel G-protein-coupled receptor up-regulated
 CC in prostate cancer, termed PHOR-1. The encoding cDNA is contained in
 CC plasmid designated p101P311 deposited with ATCC as Accession No. PTA-312.
 CC PHOR-1 polypeptides and polynucleotides are useful for diagnosing the
 CC presence of cancer, especially prostate, kidney, uterine, cervical,
 CC stomach or rectal cancer by determining and comparing the level of the
 CC protein or mRNA expression in test and normal tissue samples.

CC Pharmaceutical compositions comprising PHOR-1 is useful for treating
 CC cancer. PHOR-1 proteins are useful for identifying ligands and other
 CC agents and cellular constituents that binds to PHOR-1 gene product and
 CC for generating antibodies which are useful in diagnostic, prognostic and
 CC imaging methodologies and for the treatment of prostate cancer. Cell
 CC lines expressing PHOR-1 are useful for identifying protein-protein
 CC interactions mediated by PHOR-1. The present sequence represents the
 CC amino acid sequence of A1138218, a PHOR-1 family member

XX Sequence 167 AA;

Query Match

Best Local Similarity 100.0%; Score 25; DB 4; Length 167;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
 DB 58 SSVMH 62
 |||||

RESULT 15

ABR43245
 ID ABR43245 standard; protein; 248 AA.

XX
 AC ABR43245;

DT 07-JUL-2003 (first entry)

DE Human PMMM-6 protein SEQ ID NO:6.

XX Human; protein modification and maintenance molecule; PMMM; cytostatic;
 KW antiarteriosclerotic; anticonvulsant; nootropic; neuroprotective; AIDS;
 KW cerebroprotective; anti-HIV; antiallergic; antiinflammatory; cancer;
 KW thymimetic; gene therapy; cell proliferative disorder; atherosclerosis;
 KW neurological disorder; epilepsy; Huntington's disease; stroke; allergy;
 KW immune disorder; inflammatory disorder; developmental disorder;
 KW hypothyroidism; Cushing's syndrome; infection.

XX Homo sapiens.

XX WO2003025131-A2.

XX 27-MAR-2003.

XX 13-SEP-2002; 2002WO-US029221.

XX 14-SEP-2001; 2001US-0322196P.

XX 21-SEP-2001; 2001US-0324134P.

XX 26-OCT-2001; 2001US-0327233P.

XX 02-NOV-2001; 2001US-0343980P.

XX 09-NOV-2001; 2001US-0348887P.

XX 16-NOV-2001; 2001US-0332423P.

XX 28-NOV-2001; 2001US-0334145P.

XX 28-NOV-2001; 2001US-0334229P.

XX 06-DEC-2001; 2001US-0337451P.

XX 25-JAN-2002; 2002US-0351928P.

XX 21-MAR-2002; 2002US-0366837P.

XX (INCY-) INCYTE GENOMICS INC.

XX Sprague WW, Chawla NK, Warren BA, Tang YT, Elliott VS;

XX Marquis JP, Li JX, Griffin JA, Gietzen KJ, Yang J, Lu DAM;

XX Emerling BM, Duggan BM, Richardson TW, Lee SY, Ramkumar J, Becha SD;

XX Lehr-Mason PM, Swarnakar A, Tran UK, Kable AE, Hafalia AJA, Khare R;

XX WPI; 2003-354597/33.

XX N-PSDB; ACC59964.

XX New human protein modification and maintenance molecules (PMMM), useful
 PT for diagnosing, treating and preventing diseases or conditions associated
 PT with the aberrant PMMM expression e.g. cancer, AIDS, epilepsy, or
 PT infections.

XX Claim 1; Page 213-214; 270pp; English.

XX ACC59959 to ACC59989 encode the human protein modification and
 CC maintenance molecule proteins given in ABR43240 to ABR43270, designated
 CC PMMM-1 to PMMM-31 (I). (I) have cytostatic, antiarteriosclerotic,
 CC anticonvulsant, nootropic, neuroprotective, cerebroprotective, anti-HIV,
 CC antiinflammatory, antiinflammatory and thymimetic activities, and can be
 CC used in gene therapy. The PMMM polypeptides and polynucleotides are
 CC useful in diagnosing, treating and preventing diseases or conditions
 CC associated with the decreased expression or overexpression of PMMM, such
 CC as cell proliferative (e.g. cancer, atherosclerosis), neurological (e.g.
 CC epilepsy, Huntington's disease, stroke), immune/inflammatory (e.g. AIDS,
 CC allergies) and developmental (e.g. hypothyroidism, Cushing's syndrome)
 CC disorders, or infections. They are also useful in assessing the effects

CC of exogenous compounds on the expression of nucleic acid and amino acid
CC sequences of PMM. The PMMs or their fragments are useful in screening
CC compounds for effectiveness as agonist or antagonist of the polypeptides,
CC or in altering the expression of the target polynucleotide and compounds
CC that specifically bind to or modulate the activity of the polypeptide.
CC The microarray is useful in monitoring or measuring protein-protein
CC interactions, drug-target interactions, and gene expression profiles
XX
SQ Sequence 248 AA;
Query Match 100.0%; Score 25; DB 6; Length 248;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SSMVH 5
|||
Db 234 SSMVH 238

Search completed: April 28, 2005, 18:17:27
Job time : 40.6562 secs

This Page Blank (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2005, 18:06:50 ; Search time 10.1562 Seconds
(without alignments)
36.750 Million cell updates/sec

Title: US-10-088-639A-2_COPY_158_162
Perfect score: 25
Sequence: 1 SSMVH 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	25	100.0	144	US-09-107-532A-7091	Sequence 7091, Ap
2	25	100.0	163	US-09-680-728-7	Sequence 7, Appli
3	25	100.0	163	US-10-017-066A-7	Sequence 7, Appli
4	25	100.0	269	US-09-902-540-12352	Sequence 12352, A
5	25	100.0	300	US-09-934-551-2	Sequence 2, Appli
6	25	100.0	300	US-10-224-414-2	Sequence 2, Appli
7	25	100.0	406	US-09-489-039A-12736	Sequence 12736, A
8	25	100.0	517	US-09-134-000C-4515	Sequence 4515, Ap
9	25	100.0	726	5208144-37	Patent No. 5208144
10	25	100.0	726	5208144-37	Patent No. 5208144
11	24	96.0	23	US-08-118-270-93	Sequence 93, Appl
12	24	96.0	23	PCT-US93-08528-93	Sequence 93, Appl
13	24	96.0	56	US-09-270-767-35394	Sequence 35394, A
14	24	96.0	56	US-09-270-767-50611	Sequence 50611, A
15	24	96.0	134	US-09-248-796A-24525	Sequence 24525, A
16	24	96.0	268	US-09-934-551-4	Sequence 4, Appli
17	24	96.0	268	US-10-224-414-4	Sequence 4, Appli
18	24	96.0	327	US-09-710-279-1104	Sequence 1104, Ap
19	24	96.0	630	US-09-134-001C-4615	Sequence 4615, Ap
20	22	88.0	19	US-09-441-502B-6	Sequence 6, Appli
21	22	88.0	19	US-09-441-502B-7	Sequence 7, Appli
22	22	88.0	35	US-08-118-270-232	Sequence 232, App
23	22	88.0	35	PCT-US93-08528-232	Sequence 232, App
24	22	88.0	71	US-09-621-976-5485	Sequence 5485, Ap
25	22	88.0	86	US-09-107-532A-7054	Sequence 7054, Ap
26	22	88.0	118	US-09-270-767-48353	Sequence 48353, A
27	22	88.0	136	US-09-270-767-35513	Sequence 35513, A

28	22	88.0	136	4	US-09-270-767-50730	Sequence 50730, A
29	22	88.0	165	4	US-09-270-767-39330	Sequence 39330, A
30	22	88.0	165	4	US-09-270-767-54547	Sequence 54547, A
31	22	88.0	191	4	US-10-101-464A-625	Sequence 625, App
32	22	88.0	201	4	US-09-328-352-5968	Sequence 5968, Ap
33	22	88.0	220	4	US-09-328-352-4307	Sequence 4307, Ap
34	22	88.0	243	4	US-09-902-540-14773	Sequence 14773, A
35	22	88.0	260	2	US-07-857-224B-27	Sequence 27, Appl
36	22	88.0	267	4	US-09-663-600A-190	Sequence 190, App
37	22	88.0	275	4	US-09-538-092-958	Sequence 958, App
38	22	88.0	279	4	US-09-489-039A-13210	Sequence 13210, A
39	22	88.0	284	4	US-09-248-796A-17082	Sequence 17082, A
40	22	88.0	306	4	US-09-252-991A-19132	Sequence 19132, A
41	22	88.0	316	4	US-09-134-000C-4441	Sequence 4441, Ap
42	22	88.0	318	4	US-09-252-991A-21306	Sequence 21306, A
43	22	88.0	323	4	US-09-489-039A-13907	Sequence 13907, A
44	22	88.0	339	3	US-08-688-988-2	Sequence 2, Appli
45	22	88.0	349	4	US-09-252-991A-25984	Sequence 25984, A

ALIGNMENTS

RESULT 1

US-09-107-532A-7091

; Sequence 7091, Application US/09107532A

; Patent No. 6583275

; GENERAL INFORMATION:

; APPLICANT: Lynn A Doucette-Stamm and David Bush

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS

; NUMBER OF SEQUENCES: 7310

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: GENOME THERAPEUTICS CORPORATION

; STREET: 100 Beaver Street

; CITY: Waltham

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02354

; COMPUTER READABLE FORM:

; MEDIUM TYPE: CD-ROM ISO9660

; COMPUTER: PC

; OPERATING SYSTEM: <Unknown>

; SOFTWARE: ASCII

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/107,532A

; FILING DATE: 30-Jun-1998

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/085,598

; FILING DATE: 14 May 1998

; APPLICATION NUMBER: 60/051571

; FILING DATE: July 2, 1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Ariniello, Pamela Deneke

; REGISTRATION NUMBER: 40,489

; REFERENCE/DOCKET NUMBER: GTC-012

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (781)893-5007

; TELEFAX: (781)893-8277

; INFORMATION FOR SEQ ID NO: 7091:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 144 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; HYPOTHETICAL: YES

; ORIGINAL SOURCE:

; ORGANISM: Enterococcus faecium

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (B) LOCATION 1...144

; SEQUENCE DESCRIPTION: SEQ ID NO: 7091:

US-09-107-532A-7091

Query Match 100.0%; Score 25; DB 4; Length 144;
Best Local Similarity 100.0%; Pred. No. 1.4e+02; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
|
|
|
|
|
Db 94 SSVMH 98

RESULT 2
US-09-680-728-7
; Sequence 7, Application US/09680728
; Patent No. 6750631
; GENERAL INFORMATION:
; APPLICANT: Arthur B. Raitano
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Aya Jakobovits
; APPLICANT: Mary Faris
; APPLICANT: Rene S. Hubert
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Douglas C. Safran
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTOR
; TITLE OF INVENTION: UP-REGULATED IN PROSTATE CANCER AND USES THEREOF
; FILE REFERENCE: 129,24USUI
; CURRENT APPLICATION NUMBER: US/09/680,728
; CURRENT FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/157,902
; PRIOR FILING DATE: 1999-10-05
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-680-728-7

Query Match 100.0%; Score 25; DB 4; Length 163;
Best Local Similarity 100.0%; Pred. No. 1.6e+02; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
|
|
|
|
|
Db 58 SSVMH 62

RESULT 3
US-10-017-066A-7
; Sequence 7, Application US/10017066A
; Patent No. 6838258
; GENERAL INFORMATION:
; APPLICANT: Arthur B. Raitano
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Aya Jakobovits
; APPLICANT: Mary Faris
; APPLICANT: Rene S. Hubert
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Douglas C. Safran
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTOR
; TITLE OF INVENTION: UP-REGULATED IN PROSTATE CANCER AND USES THEREOF
; FILE REFERENCE: 511582002410
; CURRENT APPLICATION NUMBER: US/10/017,066A
; CURRENT FILING DATE: 2002-05-28
; PRIOR APPLICATION NUMBER: US 09/680,728
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/157,902
; PRIOR FILING DATE: 1999-10-05
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Homo Sapiens

US-10-017-066A-7

Query Match 100.0%; Score 25; DB 4; Length 163;
Best Local Similarity 100.0%; Pred. No. 1.6e+02; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
|
|
|
|
|
Db 58 SSVMH 62

RESULT 4
US-09-902-540-12352
; Sequence 12352, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 12352
; LENGTH: 269
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-12352

Query Match 100.0%; Score 25; DB 4; Length 259;
Best Local Similarity 100.0%; Pred. No. 2.6e+02; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
|
|
|
|
|
Db 238 SSVMH 242

RESULT 5
US-09-934-551-2
; Sequence 2, Application US/09934551
; Patent No. 6461850
; GENERAL INFORMATION:
; APPLICANT: BEASLEY, Ellen M.
; APPLICANT: Li, Zhenya
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL000757
; CURRENT APPLICATION NUMBER: US/09/934,551
; CURRENT FILING DATE: 2001-08-23
; PRIOR APPLICATION NUMBER: 60/226,903
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 300
; TYPE: PRT
; ORGANISM: HUMAN
US-09-934-551-2

Query Match 100.0%; Score 25; DB 4; Length 300;
Best Local Similarity 100.0%; Pred. No. 2.9e+02; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
|
|
|
|
|
Db 233 SSVMH 237

```
RESULT 6
US-10-224-414-2
; Sequence 2, Application US/10224414
; Patent No. 6638751
; GENERAL INFORMATION:
; APPLICANT: BEASLEY, Ellen M.
; APPLICANT: Li, Zhenya
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL000757 DIV
; CURRENT APPLICATION NUMBER: US/10/224,414
; CURRENT FILING DATE: 2002-08-21
; PRIOR APPLICATION NUMBER: 60/226,903
; PRIOR FILING DATE: 2001-08-23
; PRIOR APPLICATION NUMBER: 09/934,551
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 300
; TYPE: PRT
; ORGANISM: HUMAN
US-10-224-414-2

Query Match      100.0%; Score 25; DB 4; Length 300;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SSVMH 5
      |||||
Db      233 SSVMH 237

RESULT 7
US-09-489-039A-12736
; Sequence 12736, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 12736
; LENGTH: 406
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12736

Query Match      100.0%; Score 25; DB 4; Length 406;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SSVMH 5
      |||||
Db      273 SSVMH 277

RESULT 8
US-09-134-000C-4515
; Sequence 4515, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
```

```
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4515
; LENGTH: 517
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (477)..(500)
; OTHER INFORMATION: Amino acid 477 & 500 are Xaa wherein Xaa = any amino acid.
US-09-134-000C-4515

Query Match      100.0%; Score 25; DB 4; Length 517;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SSVMH 5
      |||||
Db      467 SSVMH 471
```

```
RESULT 9
5208144-37
; Patent No. 5208144
; APPLICANT: SMITH, JOHN A.;RAYCHOWDHURY, RAKTIMA;NILES,JOHN L.
; TITLE OF INVENTION: METHOD FOR DETECTION OF HUMAN DNA
; CONTAINING THE GENE ENCODING LOW DENSITY LIPOPROTEIN RECEPTOR
; NUMBER OF SEQUENCES: 42
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/396,697
; FILING DATE: 22-AUG-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 313,682
; FILING DATE: 22-FEB-1989
; APPLICATION NUMBER: 235,211
; FILING DATE: 23-AUG-1988
; SEQ ID NO:37:
; LENGTH: 726
5208144-37
```

```
Query Match      100.0%; Score 25; DB 6; Length 726;
Best Local Similarity 100.0%; Pred. No. 6.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 SSVMH 5
      |||||
Db      82 SSVMH 86
```

```
RESULT 10
5208144-37
; Patent No. 5208144
; APPLICANT: SMITH, JOHN A.;RAYCHOWDHURY, RAKTIMA;NILES,JOHN L.
; TITLE OF INVENTION: METHOD FOR DETECTION OF HUMAN DNA
; CONTAINING THE GENE ENCODING LOW DENSITY LIPOPROTEIN RECEPTOR
; NUMBER OF SEQUENCES: 42
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/396,697
; FILING DATE: 22-AUG-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 313,682
; FILING DATE: 22-FEB-1989
; APPLICATION NUMBER: 235,211
; FILING DATE: 23-AUG-1988
; SEQ ID NO:37:
; LENGTH: 726
5208144-37
```

```
Query Match      100.0%; Score 25; DB 6; Length 726;
```

Best Local Similarity 100.0%; Pred. No. 6.7e+02; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0;

QY 1 SSVMH 5
Db 82 SSVMH 86

RESULT 11

US-08-118-270-93
; Sequence 93, Application US/08118270
; Patent No. 5508384
; GENERAL INFORMATION:
; APPLICANT: Murphy, Randall B.
; APPLICANT: Schuster, David I.
; TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
; TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
; NUMBER OF SEQUENCES: 348
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/118,270
; FILING DATE: 09-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/943,236
; FILING DATE: 10-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Townsend, Kevin G.
; REGISTRATION NUMBER: 34,033
; REFERENCE/DOCKET NUMBER: MURPHY=2A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 93:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US93-08528-93

Query Match 96.0%; Score 24; DB 1; Length 23;
Best Local Similarity 80.0%; Pred. No. 39;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
Db 13 SSIMH 17

US-08-118-270-93
; Sequence 93, Application PC/TUS9308528
; GENERAL INFORMATION:
; APPLICANT: New York University
; TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
; TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
; NUMBER OF SEQUENCES: 348
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300

Query Match 96.0%; Score 24; DB 1; Length 23;
Best Local Similarity 80.0%; Pred. No. 39;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
Db 13 SSIMH 17

RESULT 12

PCT-US93-08528-93
; Sequence 93, Application PC/TUS9308528
; GENERAL INFORMATION:
; APPLICANT: New York University
; TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
; TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
; NUMBER OF SEQUENCES: 348
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300

CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/08528
; FILING DATE: 09-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/943,236
; FILING DATE: 10-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Townsend, Kevin G.
; REGISTRATION NUMBER: 34,033
; REFERENCE/DOCKET NUMBER: MURPHY=2 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 93:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US93-08528-93

Query Match 96.0%; Score 24; DB 5; Length 23;
Best Local Similarity 80.0%; Pred. No. 39;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
Db 13 SSIMH 17

RESULT 13

US-09-270-767-35394
; Sequence 35394, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 35394
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-35394

Query Match 96.0%; Score 24; DB 4; Length 56;
Best Local Similarity 80.0%; Pred. No. 94;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
Db 24 SSIMH 28

RESULT 14

US-09-270-767-50611
; Sequence 50611, Application US/09270767

```
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 50611
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-50611
```

```
Query Match          96.0%; Score 24; DB 4; Length 56;
Best Local Similarity 80.0%; Pred. No. 94;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 SSVMH 5
        ||:||
Db      24 SSIMH 28
```

```
RESULT 15
US-09-248-796A-24525
; Sequence 24525, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 24525
; LENGTH: 134
; TYPE: PRT
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (101)
; OTHER INFORMATION: Identity of amino acid sequences at the above locations are unkno
US-09-248-796A-24525
```

```
Query Match          96.0%; Score 24; DB 4; Length 134;
Best Local Similarity 80.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 SSVMH 5
        ||:||
Db     108 SSIMH 112
```

```
Search completed: April 28, 2005, 18:28:08
Job time : 12.1562 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2005, 18:24:46 ; Search time 31.0156 Seconds
(without alignments)
53.700 Million cell updates/sec

Title: US-10-088-639a-2_COPY_158_162

Perfect score: 25

Sequence: 1 SSMVMH 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1426032 seqs, 333106140 residues

Total number of hits satisfying chosen parameters: 1426032

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	25	100.0	69	11	US-09-864-408A-5172
2	25	100.0	76	9	US-09-864-761-43019
3	25	100.0	96	15	US-10-276-774-2118
4	25	100.0	112	11	US-09-980-049-5
5	25	100.0	147	16	US-10-437-963-182859
6	25	100.0	252	11	US-10-145-586-37
7	25	100.0	253	11	US-09-981-566A-51
8	25	100.0	261	14	US-10-017-161-966
9	25	100.0	282	14	US-10-167-555-2
10	25	100.0	300	9	US-09-934-551-2
11	25	100.0	300	14	US-10-224-414-2
12	25	100.0	300	14	US-10-143-575-2
13	25	100.0	300	15	US-10-640-326-2

Sequence 24, Appl
Sequence 85, Appl
Sequence 89, Appl
Sequence 85, Appl
Sequence 89, Appl
Sequence 26, Appl
Sequence 10, Appl
Sequence 88, Appl
Sequence 4, Appl
Sequence 162, Appl
Sequence 164, Appl
Sequence 756, Appl
Sequence 758, Appl
Sequence 160, Appl
Sequence 162, Appl
Sequence 35, Appl
Sequence 12, Appl
Sequence 1320, Appl
Sequence 20, Appl
Sequence 22, Appl
Sequence 73027, A
Sequence 75084, A
Sequence 75831, A
Sequence 60141, A
Sequence 2, Appl
Sequence 10822, A
Sequence 182856, A
Sequence 166420, A
Sequence 74, Appl
Sequence 37219, A

306 10 US-09-844-861A-24
314 9 US-09-886-055-85
314 9 US-09-886-055-89
314 10 US-09-804-231-85
314 10 US-09-804-231-89
314 10 US-09-844-861A-26
314 13 US-10-032-106-10
314 14 US-10-017-161-878
314 14 US-10-017-161-880
314 15 US-10-182-822A-4
314 15 US-10-387-629-162
314 15 US-10-387-629-164
314 15 US-10-292-798-756
314 15 US-10-292-798-758
314 15 US-10-343-650A-160
314 15 US-10-343-650A-162
314 14 US-10-145-586-35
316 15 US-10-466-720-12
316 17 US-10-774-355A-1320
317 10 US-09-844-861A-22
322 10 US-09-844-861A-22
352 15 US-10-282-122A-75084
352 15 US-10-282-122A-75084
352 15 US-10-282-122A-75831
353 15 US-10-282-122A-60141
436 14 US-10-078-592-2
491 15 US-10-369-493-10822
728 16 US-10-437-963-182856
1293 16 US-10-437-963-166420
4660 15 US-10-464-368-74
58 9 US-09-864-761-37219

ALIGNMENTS

RESULT 1

US-09-864-408A-5172
; Sequence 5172, Application US/09864408A
; Publication No. US20040009474A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Shimketa, Richard A.
; TITLE OF INVENTION: No. US20040009474A1el Human Polynucleotides and Polypeptides Encod
; FILE REFERENCE: 21402-012
; CURRENT APPLICATION NUMBER: US/09/864,408A
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/206,690
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 9068
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5172
; LENGTH: 69
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-864-408A-5172

Query Match 100.0%; Score 25; DB 11; Length 69;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSMVMH 5

Db 23 SSMVMH 27

RESULT 2

US-09-864-761-43019
; Sequence 43019, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.

APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
FILE REFERENCE: Aemica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 43019
LENGTH: 76
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC009758.1
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.92
OTHER INFORMATION: SWISSPROT HIT: O95156, EVALUATE 6.00e-05
OTHER INFORMATION: EST_HUMAN HIT: A1138213.1, EVALUATE 4.00e-25
NAME/KEY: unsure
LOCATION: 13
US-09-864-761-43019

Query Match 100.0%; Score 25; DB 9; Length 76;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
Db 35 SSVMH 39
|||||

RESULT 3
US-10-276-774-2118
Sequence 2118, Application US/10276774

Publication No. US20040053245A1
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
APPLICANT: Tang, Y, Tom et al
TITLE OF INVENTION: No. US20040053245A1el Nucleic Acids and Polypeptides
FILE REFERENCE: 21272-030
CURRENT APPLICATION NUMBER: US/10/276,774
CURRENT FILING DATE: 2002-11-18
PRIOR APPLICATION NUMBER: 09/560,875
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 09/496,914
PRIOR FILING DATE: 2000-02-03
NUMBER OF SEQ ID NOS: 2700
SOFTWARE: Custom
SEQ ID NO 2118
LENGTH: 96
TYPE: PRT
ORGANISM: Homo sapiens
US-10-276-774-2118

Query Match 100.0%; Score 25; DB 15; Length 96;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
Db 39 SSVMH 43
|||||

RESULT 4

US-09-980-049-5
Sequence 5, Application US/09980049
Publication No. US20040220092A1
GENERAL INFORMATION:
APPLICANT: INCYTE GENOMICS, INC.
APPLICANT: POLICKY, Jennifer L.
APPLICANT: TRIBOULEY, Catherine M.
APPLICANT: TANG, Y. Tom
APPLICANT: BAUGHN, Mariah R.
APPLICANT: GRAUL, Richard
APPLICANT: KHAN, Farrah A.
APPLICANT: NGUYEN, Dannel B.
APPLICANT: PATTERSON, Chandra
APPLICANT: LAL, Preeti
APPLICANT: AU-YOUNG, Janice
APPLICANT: YANG, Junming
APPLICANT: HAFALIA, April
APPLICANT: WALIA, Narinder K.
APPLICANT: DAS, Debopriya
TITLE OF INVENTION: G-PROTEIN COUPLED RECEPTORS
FILE REFERENCE: PI-0072 PCT
CURRENT APPLICATION NUMBER: US/09/980,049
CURRENT FILING DATE: 2001-11-28
PRIOR APPLICATION NUMBER: 60/193,051; 60/195,155; 60/199,084; 60/200,551; 60/202,278
PRIOR FILING DATE: 2000-03-29; 2000-04-06; 2000-04-20; 2000-04-28; 2000-05-05
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PERL Program
SEQ ID NO 5
LENGTH: 112
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No: 4585651CD1
US-09-980-049-5

Query Match 100.0%; Score 25; DB 11; Length 112;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
Db 55 SSVMH 59
|||||

RESULT 5
US-10-437-963-182859
; Sequence 182859, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 182859
; LENGTH: 147
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(147)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_80004C.1.pap
US-10-437-963-182859

Query Match 100.0%; Score 25; DB 16; Length 147;
Best Local Similarity 100.0%; Pred. No. 3.3e+02; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0;

QY 1 SSVMH 5
|
|
|
|
|
Db 60 SSVMH 64

RESULT 6
US-10-145-586-37
; Sequence 37, Application US/10145586
; Publication No. US2003013890A1
; GENERAL INFORMATION:
; APPLICANT: Alexandra Glucksmann, Maria
; APPLICANT: Silos-Santiago, Inmaculada
; APPLICANT: M. Galvin, Katherine
; APPLICANT: Weich, Nadine
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Bandaru, Rajasekhara
; APPLICANT: Kapeller-Libermann, Rosana.
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTOR FAMILY MEMBERS,
; TITLE OF INVENTION: HUMAN THIOREDOXIN FAMILY MEMBERS, HUMAN LEUCINE-RICH
; TITLE OF INVENTION: REPEAT FAMILY MEMBERS, AND HUMAN RINGFINGER FAMILY MEMBER
; FILE REFERENCE: 10448-188001
; CURRENT APPLICATION NUMBER: US/10/145,586
; CURRENT FILING DATE: 2002-05-14
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-145-586-37

Query Match 100.0%; Score 25; DB 14; Length 252;
Best Local Similarity 100.0%; Pred. No. 5.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
|
|
|
|
|
Db 218 SSVMH 222

RESULT 7

US-09-981-566A-51
; Sequence 51, Application US/09981566A
; Publication No. US20040005656A1
; GENERAL INFORMATION:
; APPLICANT: Kekuda et al.
; TITLE OF INVENTION: No. US20040005656A1el GPCR-like Proteins and Nucleic Acids Encod
; FILE REFERENCE: 21402-163
; CURRENT APPLICATION NUMBER: US/09/981,566A
; CURRENT FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 60/240,704
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/262,159
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/263,340
; PRIOR FILING DATE: 2001-01-22
; PRIOR APPLICATION NUMBER: 60/264,118
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/308,203
; PRIOR FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: 60/243,497
; PRIOR FILING DATE: 2000-10-26
; PRIOR APPLICATION NUMBER: 60/244,542
; PRIOR FILING DATE: 2000-10-31
; PRIOR APPLICATION NUMBER: 60/269,031
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: 60/245,484
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/255,017
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: 60/263,216
; PRIOR FILING DATE: 2001-01-22
; PRIOR APPLICATION NUMBER: 60/268,225
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 209
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 51
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-981-566A-51

Query Match 100.0%; Score 25; DB 11; Length 253;
Best Local Similarity 100.0%; Pred. No. 5.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
|
|
|
|
|
Db 199 SSVMH 203

RESULT 8

US-10-017-161-966
; Sequence 966, Application US/10017161
; Publication No. US20030143668A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 08435/0152
; CURRENT APPLICATION NUMBER: US/10/017,161
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP 2001/246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2430

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 966
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-017-161-966

Query Match      100.0%; Score 25; DB 14; Length 261;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 SSVMH 5
Db      204 SSVMH 208
      |||||
      |||||

RESULT 9
US-10-167-555-2
; Sequence 2, Application US/10167555
; Publication No. US20030022212A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
; TITLE OF INVENTION: 65649, A Human Metalloprotease Family Member and Uses
; TITLE OF INVENTION: Therefor
; FILE REFERENCE: MP101-089P1RM
; CURRENT APPLICATION NUMBER: US/10/167,555
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: 60/297,938
; PRIOR FILING DATE: 2001-06-13
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 282
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-167-555-2

Query Match      100.0%; Score 25; DB 14; Length 282;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 SSVMH 5
Db      233 SSVMH 237
      |||||
      |||||

RESULT 10
US-09-934-551-2
; Sequence 2, Application US/09934551
; Publication No. US20020072106A1
; GENERAL INFORMATION:
; APPLICANT: BEASLEY, Ellen M.
; APPLICANT: Li, Zhenya
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL000757
; CURRENT APPLICATION NUMBER: US/09/934,551
; CURRENT FILING DATE: 2001-08-23
; PRIOR APPLICATION NUMBER: 60/226,903
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 300
; TYPE: PRT
; ORGANISM: HUMAN
US-09-934-551-2

Query Match      100.0%; Score 25; DB 9; Length 300;
Best Local Similarity 100.0%; Pred. No. 6.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 SSVMH 5
Db      233 SSVMH 237
      |||||
      |||||

RESULT 11
US-10-224-414-2
; Sequence 2, Application US/10224414
; Publication No. US20030036167A1
; GENERAL INFORMATION:
; APPLICANT: BEASLEY, Ellen M.
; APPLICANT: Li, Zhenya
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL000757 DIV
; CURRENT APPLICATION NUMBER: US/10/224,414
; CURRENT FILING DATE: 2002-08-21
; PRIOR APPLICATION NUMBER: 60/226,903
; PRIOR FILING DATE: 2001-08-23
; PRIOR APPLICATION NUMBER: 09/934,551
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 300
; TYPE: PRT
; ORGANISM: HUMAN
US-10-224-414-2

Query Match      100.0%; Score 25; DB 14; Length 300;
Best Local Similarity 100.0%; Pred. No. 6.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 SSVMH 5
Db      233 SSVMH 237
      |||||
      |||||

RESULT 12
US-10-143-575-2
; Sequence 2, Application US/10143575
; Publication No. US20030166072A1
; GENERAL INFORMATION:
; APPLICANT: ZHONG, Jenny M. et al.
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL001240
; CURRENT APPLICATION NUMBER: US/10/143,575
; CURRENT FILING DATE: 2002-05-13
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-143-575-2

Query Match      100.0%; Score 25; DB 14; Length 300;
Best Local Similarity 100.0%; Pred. No. 6.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 SSVMH 5
Db      233 SSVMH 237
      |||||
      |||||

RESULT 13
US-10-640-326-2
; Sequence 2, Application US/10640326
; Publication No. US20040038896A1
; GENERAL INFORMATION:
```

; APPLICANT: BEASLEY, Ellen M.
; APPLICANT: LI, Zhenya
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CLO00757 DIV2
; CURRENT APPLICATION NUMBER: US/10/640,326
; PRIOR FILING DATE: 2003-08-25
; PRIOR APPLICATION NUMBER: 60/226,903
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: 09/934,551
; PRIOR FILING DATE: 2001-08-23
; PRIOR APPLICATION NUMBER: 10/224,414
; PRIOR FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 300
; TYPE: PRT
; ORGANISM: HUMAN
US-10-640-326-2

Query Match 100.0%; Score 25; DB 15; Length 300;
Best Local Similarity 100.0%; Pred. No. 6.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
|||||
Db 233 SSVMH 237

RESULT 14

US-09-844-861A-24
; Sequence 24, Application US/09844861A
; Publication No. US20030216304A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Mishra, Vishnu
; APPLICANT: Spytek, Kimberly
; APPLICANT: Burgess, Catherine
; APPLICANT: Lepley, Denise
; APPLICANT: Grosse, William
; APPLICANT: Szekeres, Edward
; APPLICANT: Alsobrook, John
; APPLICANT: Gangolli, Esha
; APPLICANT: Casman, Stacie
; APPLICANT: MacDougall, John
; APPLICANT: Smithson, Glennda
; TITLE OF INVENTION: No. US20030216304A1 Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-789 US
; CURRENT APPLICATION NUMBER: US/09/844,861A
; CURRENT FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: 60/199,947
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/199,960
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/225,226
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/256,399
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/256,524
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/258,159
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/258,511
; PRIOR FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: 60/258,828
; PRIOR FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: 60/259,659
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/275,604
; PRIOR FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 113

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 24
; LENGTH: 306
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-844-861A-24

Query Match 100.0%; Score 25; DB 10; Length 306;
Best Local Similarity 100.0%; Pred. No. 6.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
|||||
Db 249 SSVMH 253

RESULT 15

US-09-886-055-85
; Sequence 85, Application US/09886055
; Patent No. US2002013273A1
; GENERAL INFORMATION:
; APPLICANT: STRYER, LUBERT
; APPLICANT: ZOZULYA, SERGEY
; TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND
; TITLE OF INVENTION: BIOSENSORS OF CHEMICAL SENSANTS
; FILE REFERENCE: 078003-0277150
; CURRENT APPLICATION NUMBER: US/09/886,055
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/213,812
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 522
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 85
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-886-055-85

Query Match 100.0%; Score 25; DB 9; Length 314;
Best Local Similarity 100.0%; Pred. No. 7.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
|||||
Db 260 SSVMH 264

Search completed: April 28, 2005, 19:02:16
Job time : 33.0156 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 01:42:46 ; Search time 4.15094 Seconds
(without alignments)
62.721 Million cell updates/sec

Title: US-10-088-639A-2_COPY_158_162
Perfect score: 25
Sequence: 1 SSVMH 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	25	100.0	244	1	CAH_ERWCA
2	25	100.0	310	1	ASTL_COTJA
3	25	100.0	314	1	OYK1_HUMAN
4	25	100.0	314	1	OYK2_HUMAN
5	25	100.0	352	1	MODC_SALTI
6	25	100.0	352	1	MODC_SALTY
7	25	100.0	353	1	HMZ1_DROME
8	25	100.0	373	1	HMZ1_DROSU
9	25	100.0	379	1	CYB_CRATY
10	25	100.0	379	1	CYB_MESPE
11	25	100.0	380	1	CYB_PERTA
12	25	100.0	380	1	CYB_PUNCO
13	25	100.0	380	1	CYB_STRPU
14	25	100.0	387	1	NMT_DROME
15	25	100.0	396	1	CYB_LAMFL
16	25	100.0	396	1	CYB_PETMA
17	25	100.0	416	1	RHLA_PANTR
18	25	100.0	416	1	RHLG_GORGO
19	25	100.0	416	1	RHLJ_GORGO
20	25	100.0	416	1	RHLF_PANTR
21	25	100.0	1075	1	NFC3_MOUSE
22	25	100.0	4660	1	LRP2_RAT
23	24	96.0	90	1	RPOZ_VIBCH
24	24	96.0	128	1	CYB_GROVV
25	24	96.0	214	1	CYB_AFRMI
26	24	96.0	214	1	CYB_LACMU
27	24	96.0	270	1	HCE1_ORYLA
28	24	96.0	279	1	HCE2_ORYLA
29	24	96.0	330	1	YD31_ARCTU
30	24	96.0	371	1	CYB_ERYCC
31	24	96.0	892	1	VP48_VARV
32	22	88.0	92	1	YC48_STREY
33	22	88.0	109	1	RPCH_CALSI
					052538 erwinia car
					P42662 coturnix co
					Q8ngk4 homo sapien
					Q8ngk3 homo sapien
					Q828a4 salmonella
					Q82qr6 salmonella
					P09089 drosophila
					Q24648 drosophila
					Q8wdul cratogeomys
					Q94t21 percopsis t
					Q922c4 puntius con
					P15547 strongyloce
					Q61613 drosophila
					Q9ta00 lampetra fl
					Q35534 petromyzon
					Q28813 pan troglod
					Q28426 gorilla gor
					Q28427 gorilla gor
					P28812 pan troglod
					P97305 mus musculus
					P98158 rattus norv
					Q9kxm3 vibrio chol
					P95776 crotalus vi
					P87421 atractaspis
					P92853 lachesis mu
					P31580 oryzias lat
					P31581 oryzias lat
					Q28938 archaeoglob
					O48066 eryx colubr
					P3817 variola vir
					Q992f8 streptococc
					Q23757 callinectes

34	22	88.0	110	1	RPCH_CARMA
35	22	88.0	133	1	GEPE_BACSU
36	22	88.0	154	1	Y393_BUCBP
37	22	88.0	173	1	CD3D_MOUSE
38	22	88.0	180	1	YOM6_CABEL
39	22	88.0	198	1	YCF4_ANASP
40	22	88.0	211	1	TRMB_STRMU
41	22	88.0	275	1	RPB3_HUMAN
42	22	88.0	275	1	RPB3_MOUSE
43	22	88.0	284	1	YNFH_ECOLI
44	22	88.0	301	1	VG30_BPPF3
45	22	88.0	312	1	GU27_RAT
46	22	88.0	320	1	O7C2_HUMAN
47	22	88.0	332	1	AAIP_WHEAT
48	22	88.0	342	1	TAF5_HUMAN
49	22	88.0	346	1	RECA_CHLTE
50	22	88.0	355	1	YP7I_MYCTU
51	22	88.0	375	1	ANM6_HUMAN
52	22	88.0	385	1	HEM6_SOYBN
53	22	88.0	386	1	HEM6_ARATH
54	22	88.0	391	1	GAT5_CHICK
55	22	88.0	391	1	HEM6_HORVU
56	22	88.0	397	1	HEM6_TOBAC
57	22	88.0	402	1	FIXF_RHISN
58	22	88.0	428	1	FXB2_MOUSE
59	22	88.0	450	1	SYN_MYCPU
60	22	88.0	451	1	GPID_CHLTR
61	22	88.0	466	1	DNAA_DEIRA
62	22	88.0	475	1	LEGU_CANEN
63	22	88.0	481	1	ALYS_STAAR
64	22	88.0	490	1	CPC2_RABIT
65	22	88.0	490	1	CPC6_RAT
66	22	88.0	490	1	CRCP_MESAU
67	22	88.0	490	1	CRCP_MESAU
68	22	88.0	490	1	CPCR_MESAU
69	22	88.0	490	1	CPCT_MOUSE
70	22	88.0	490	1	CPZ2_MOUSE
71	22	88.0	492	1	TRPE_PSEAE
72	22	88.0	493	1	TRPE_PSEAV
73	22	88.0	498	1	NMT2_BOVIN
74	22	88.0	498	1	NMT2_HUMAN
75	22	88.0	500	1	TRPE_HELPF
76	22	88.0	500	1	TRPE_HELPY
77	22	88.0	502	1	RK11_SECCCE
78	22	88.0	505	1	TRPE_PSESS
79	22	88.0	514	1	UN52_XENLA
80	22	88.0	520	1	YMT2_MYCTU
81	22	88.0	529	1	NMT2_MOUSE
82	22	88.0	548	1	TCPO_HUMAN
83	22	88.0	548	1	TCPO_MOUSE
84	22	88.0	563	1	LIP2_GEOGN
85	22	88.0	565	1	PHR_YEAST
86	22	88.0	576	1	NU5M_ANOQU
87	22	88.0	606	1	Z214_HUMAN
88	22	88.0	649	1	SCAG_HUMAN
89	22	88.0	650	1	SCAG_MOUSE
90	22	88.0	655	1	MEPB_HUMAN
91	22	88.0	700	1	MEPB_MOUSE
92	22	88.0	704	1	MEPB_RAT
93	22	88.0	717	1	P5C3_ACTCH
94	22	88.0	738	1	GCP4_ARATH
95	22	88.0	739	1	GCP4_MEDTR
96	22	88.0	752	1	K6AA_CHICK
97	22	88.0	757	1	ECR_FUCCU
98	22	88.0	777	1	UNC8_CABEL
99	22	88.0	786	1	K6PP_RAT
100	22	88.0	802	1	Y250_HUMAN
101	22	88.0	827	1	K685_HUMAN
102	22	88.0	1064	1	KINI_YEAST
103	22	88.0	1084	1	DPYD_CABEL
104	22	88.0	1084	1	DPYD_CABEL
105	22	88.0	1159	1	NI24_SCHPO
106	22	88.0	1181	1	HAIR_RAT

Q26324	carcinus ma
O06717	bacillus eu
P59478	buchnera ap
P04235	mus musculus
P30852	caenorhabdi
Q8ypa9	anabaena sp
Q8dvq4	streptococc
P19387	homo sapien
P77160	mus musculus
P76173	escherichia
P03626	bacterioph
P34987	rattus norv
O60412	homo sapien
Q02066	triticum ae
Q06944	homo sapien
O52393	chlorobium
Q50650	mycobacteri
Q961a8	homo sapien
P35055	glycine max
Q91r75	arabidopsis
P43692	gallus gall
Q42840	hordeum vul
Q42946	nicotiana t
P55467	rhizobium s
Q64733	mus musculus
Q98pf2	mycoplasma
P10555	chlamydia t
Q9rye7	deinococcus
P49046	canavalia e
P24556	staphylococ
P00181	oryctolagus
P05178	rattus norv
Q08078	mesocricetu
P32663	mesocricetu
P32664	mesocricetu
P56654	mus musculus
P20580	pseudomonas
P20579	pseudomonas
Q9n181	bos taurus
O60551	homo sapien
Q9zju5	helicobacte
O25869	helicobacte
Q02723	secale cere
P21689	pseudomonas
P42664	xenopus lae
Q10508	mycobacteri
O70311	mus musculus
P50990	homo sapien
P42932	mus musculus
P22394	geotrichum
P05066	saccharomyc
P33510	anopheles q
Q9ul59	homo sapien
P51170	homo sapien
P37091	rattus norv
Q9wu39	mus musculus
Q16820	homo sapien
Q61847	mus musculus
P28826	rattus norv
O04015	a delta 1-p
Q9n350	arabidopsis
Q9sc88	medicago tr
P18652	gallus gall
O18531	lucilia cup
Q21974	caenorhabdi
P47860	rattus norv
Q92540	homo sapien
O75170	homo sapien
P3185	saccharomyc
O18164	caenorhabdi
Q09904	schizosacch
P97609	rattus norv

```

107 22 88.0 1182 1 HAIR_MOUSE
108 22 88.0 1204 1 DNBI_VZVD
109 22 88.0 1265 1 SAI_XENLA
110 22 88.0 1302 1 MDRA_DROME
111 22 88.0 1397 1 CID_DROME
112 22 88.0 1422 1 PTFG_CHICK
113 22 88.0 1442 1 PTFG_MOUSE
114 22 88.0 1445 1 PTFG_HUMAN
115 22 88.0 1581 1 ACC8_CRICR
116 22 88.0 1581 1 ACC8_RAT
117 22 88.0 1812 1 SUIS_SUNMU
118 22 88.0 1826 1 SUIS_RABIT
119 22 88.0 2035 1 M18A_MOUSE
120 22 88.0 2054 1 M18A_MOUSE
121 22 88.0 2104 1 MYB3_SCHPO
122 22 88.0 2452 1 RPB1_PLAFD
123 22 88.0 3172 1 ERV3_SACER
124 22 88.0 3255 1 POLG_LMV0
125 22 88.0 3255 1 POLG_LMVE
126 21 84.0 60 1 SH2B_FIG
127 21 84.0 72 1 YCX1_DENBE
128 21 84.0 110 1 VL3_HPV5B
129 21 84.0 119 1 YEAR_ECOLI
130 21 84.0 142 1 USPC_ECOLI
131 21 84.0 144 1 PCOE_ECOLI
132 21 84.0 146 1 HBB_CARCR
133 21 84.0 146 1 SP22_BACSH
134 21 84.0 147 1 HBBA_ONCMY
135 21 84.0 147 1 HBBA_ANGAN
136 21 84.0 147 1 HBB_CARAU
137 21 84.0 147 1 HBB_CVPCA
138 21 84.0 147 1 HBB_ONCNE
139 21 84.0 147 1 HBB_SALSA
140 21 84.0 158 1 Y488_METJA
141 21 84.0 174 1 NEUU_HUMAN
142 21 84.0 174 1 NEUU_MOUSE
143 21 84.0 174 1 NEUU_RAT
144 21 84.0 174 1 YIF3_PSEAE
145 21 84.0 188 1 SH2A_CANFA
146 21 84.0 188 1 C561_ECOLI
147 21 84.0 189 1 YMS0_WCTU
148 21 84.0 194 1 MUCA_PSEAE
149 21 84.0 207 1 CYOC_PSEPU
150 21 84.0 214 1 CYB_ELASE

ALIGNMENTS

RESULT 1
CAH_ERWCA STANDARD; PRT; 244 AA.
AC O52538;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE Carbonic anhydrase precursor (EC 4.2.1.1) (Carbonate dehydratase).
GN CAH.
OS Erwinia carotovora.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pectobacterium.
OX NCBI_TaxID=554;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99030316; PubMed=9811652;
RA Beach M.B., Osuna R.;
RT "Identification and characterization of the fis operon in enteric
bacteria.";
RL J. Bacteriol. 180:5932-5946(1998).
CC -!- FUNCTION: Reversible hydration of carbon dioxide.
CC -!- CATALYTIC ACTIVITY: H(2)CO(3) = CO(2) + H(2)O.
CC -!- COFACTOR: Zinc (By similarity).
CC -!- SUBCELLULAR LOCATION: Periplasmic (Potential).

```

Q61645 mus musculus
P03246 varicella-z
Q09dnl xenopus lae
Q00449 drosophila
P19538 drosophila
Q98936 gallus gall
Q05909 mus musculus
P23470 homo sapien
Q09427 cricetus cr
Q09429 rattus norv
Q62653 suncus muri
P07768 oryctolagus
Q91mh9 mus musculus
Q92614 homo sapien
O14157 schizosacch
P14248 plasmodium
Q03133 saccharopol
P31999 l genome po
P89876 l genome po
Q29005 sus scrofa
P11673 oenothera b
P26541 human papil
P76248 escherichia
P46888 escherichia
Q47459 escherichia
Q10733 carretta car
Q32724 bacillus sp
P02141 oncorhynchu
P80946 anguilla an
P02140 carassius a
P02139 cyprinus ca
Q98c80 oncorhynchu
Q91473 salmo salar
P48645 homo sapien
Q9gxk8 mus musculus
P12760 rattus norv
P40882 pseudomonas
Q46635 canis famli
P75925 escherichia
Q10528 mycobacteri
P38107 pseudomonas
Q9wvr3 pseudomonas
P87416 elapsoidea

CC -!- SIMILARITY: Belongs to the eukaryotic-type carbonic anhydrase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: AF040381; AAC77891.1; -
CC HSP: Q50940; 1KQO.
CC InterPro: IPR001148; Euk Coanhd.
CC Pfam: PF00194; carb. anhydrase; 1.
CC ProDom: PD000865; Euk Coanhd; 1.
CC PROSITE: PS00162; Euk CO2 ANHYDRASE; 1.
CC Lyase; Zinc; Periplasmic; Signal.
CC SIGNAL 1 19 POTENTIAL.
CC CHAIN 20 244 CARBONIC ANHYDRASE.
CC METAL 109 109 ZINC (CATALYTIC) (BY SIMILARITY).
CC METAL 111 111 ZINC (CATALYTIC) (BY SIMILARITY).
CC METAL 128 128 ZINC (CATALYTIC) (BY SIMILARITY).
CC DISULFID 46 199 BY SIMILARITY.
CC SEQUENCE 244 AA; 26773 MW; DCA6C1DCB9C97F55 CRC64;

Query Match 100.0%; Score 25; DB 1; Length 244;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SSVMH 5
Db 223 SSVMH 227

RESULT 2
ASTL_COTJA STANDARD; PRT; 310 AA.
ID ASTL_COTJA
AC P42662;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Astacin like metalloendopeptidase (EC 3.4.24.-).
DE Coturnix coturnix japonica (Japanese quail).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
OC Coturnix.
OX NCBI_TaxID=93934;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Chorioallantoic membrane;
RX MEDLINE=94114563; PubMed=8286408;
RA Elarousi M.A., Deluca H.F.;
RT "A new member to the astacin family of metalloendopeptidases: a novel
1,25-dihydroxyvitamin D-3-stimulated mRNA from chorioallantoic
membrane of quail.";
RL Biochim. Biophys. Acta 1217:1-8(1994).
CC -!- FUNCTION: May play a role in the degradation of eggshell matrix.
CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -!- INDUCTION: By 1,25-dihydroxyvitamin D-3.
CC -!- SIMILARITY: Belongs to peptidase family M12A.
CC -!- SIMILARITY: Contains 1 CUB domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: U12642; AAA20842.1; -
CC PIR: S41055; S41055.

```

DR HSP; P07584; 11AE.
DR MEROPS; M12.0PA; -.
DR InterPro; IPR000859; CUB.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR006026; PeptIdase_M.
DR InterPro; IPR001506; Peptidase_M12A.
DR Pfam; PF01400; Ascacin; 1.
DR Pfam; PF00431; CUB; 1.
DR PRINTS; PR00480; ASTACIN.
DR SMART; SM00042; CUB; 1.
DR SMART; SM00235; ZnMc; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; FALSE_NEG.
DR PROSITE; PS01180; CUB; 1.
KW Hydrolase; Metalloprotease; zinc.
FT METAL 84 84 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT SITE 85 85 BY SIMILARITY.
FT METAL 88 88 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 94 94 ZINC (CATALYTIC) (BY SIMILARITY).
FT DOMAIN 188 300 CUB.
SQ SEQUENCE 310 AA; 35144 MW; E2ACE1EAC96366A1 CRC64;

Query Match 100.0%; Score 25; DB 1; Length 310;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
DB 136 SSVMH 140

RESULT 3
ID OYK1 HUMAN STANDARD; PRT; 314 AA.
AC Q8NGK4;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Olfactory receptor 52K1.
GN OR52K1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Suwa M., Sato T., Okouchi I., Arita M., Futami K., Matsumoto S.,
RA Tautsumi S., Aburatani H., Asai K., Akiyama Y.;
RT "Genome-wide discovery and analysis of human seven transmembrane helix
RT receptor genes.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
CC -!- FUNCTION: Putative odorant receptor.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC -!- DATABASE: NAME=Human Olfactory Receptor Data Exploratorium (HORDE);
WWW="http://bioinformatics.weizmann.ac.il/cgi-bin/HORDE/showGene.pl?key=symbols&
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; AB065790; BAC06009.1; -.
DR Genew; HGNC:15222; OR52K1.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
Multigene family; Olfaction.
FT DOMAIN 1 27 EXTRACELLULAR (POTENTIAL).

```

```

FT TRANSMEM 28 48
FT DOMAIN 49 56 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 57 77 2 (POTENTIAL).
FT DOMAIN 78 101 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 102 122 3 (POTENTIAL).
FT DOMAIN 123 141 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 142 162 4 (POTENTIAL).
FT DOMAIN 163 198 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 199 219 5 (POTENTIAL).
FT DOMAIN 220 239 6 (POTENTIAL).
FT TRANSMEM 240 260 7 (POTENTIAL).
FT DOMAIN 261 275 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 276 296 314 CYTOPLASMIC (POTENTIAL).
FT DISULFID 99 191 BY SIMILARITY.
FT CARBOHYD 5 N-LINKED (GLCNAC... ) (POTENTIAL).
SQ SEQUENCE 314 AA; 35289 MW; 8045963557FAB634 CRC64;

Query Match 100.0%; Score 25; DB 1; Length 314;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
DB 260 SSVMH 264

RESULT 4
ID OYK2 HUMAN STANDARD; PRT; 314 AA.
AC Q8NGK3;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Olfactory receptor 52K2.
GN OR52K2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Suwa M., Sato T., Okouchi I., Arita M., Futami K., Matsumoto S.,
RA Tautsumi S., Aburatani H., Asai K., Akiyama Y.;
RT "Genome-wide discovery and analysis of human seven transmembrane helix
RT receptor genes.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
CC -!- FUNCTION: Putative odorant receptor.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC -!- DATABASE: NAME=Human Olfactory Receptor Data Exploratorium (HORDE);
WWW="http://bioinformatics.weizmann.ac.il/cgi-bin/HORDE/showGene.pl?key=symbols&
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; AB065791; BAC06010.1; -.
DR Genew; HGNC:15223; OR52K2.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor; Transmembrane.
FT DOMAIN 1 27 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 28 48 1 (POTENTIAL).
FT DOMAIN 49 56 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 57 77 2 (POTENTIAL).
FT DOMAIN 78 101 EXTRACELLULAR (POTENTIAL).

```

FT TRANSMEM 102 122 3 (POTENTIAL).
 FT DOMAIN 123 141 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 142 162 4 (POTENTIAL).
 FT DOMAIN 163 198 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 199 219 5 (POTENTIAL).
 FT DOMAIN 220 239 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 240 260 6 (POTENTIAL).
 FT DOMAIN 261 275 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 276 296 7 (POTENTIAL).
 FT DOMAIN 297 314 CYTOPLASMIC (POTENTIAL).
 FT DISULFID 99 191 BY SIMILARITY.
 FT CARBOHYD 5 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 314 AA; 34907 MW; 56D56A9898904BC58 CRC64;

Query Match 100.0%; Score 25; DB 1; Length 314;
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
 Db 260 SSVMH 264
 |||||

RESULT 5
 MODC_SALTY STANDARD; PRT; 352 AA.
 AC Q828A4;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Molybdenum import ATP-binding protein modC (EC 3.6.3.29).
 GN MODC OR STY0816 OR T2104.
 OS Salmonella typhi.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 OX NCBI_TaxID=601;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CT18;
 RX MEDLINE=21534947; PubMed=11677608;
 RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
 RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
 RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
 RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
 RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagsis K.,
 RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
 RA Quail M.A., Rutherford K., Simmonds M., Skelton J., Stevens K.,
 RA Whitehead S., Barrall B.G.;
 RT "Complete genome sequence of a multiple drug resistant Salmonella
 RT enterica serovar Typhi CT18.";
 RL Nature 413:848-852(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Ty2 / ATCC 700931;
 RX MEDLINE=22531367; PubMed=12644504;
 RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
 RA Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.;
 RT "Comparative Genomics of Salmonella enterica serovar Typhi strains Ty2
 RT and CT18.";
 RL J. Bacteriol. 185:2330-2337(2003).
 CC -!- FUNCTION: Part of the ABC transporter complex modABC (TC 3.A.1.8)
 CC involved in molybdenum import. Responsible for energy coupling to
 CC the transport system (By similarity).
 CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + molybdate (Out) = ADP + phosphate
 CC + molybdate (In).
 CC -!- SUBUNIT: The complex is composed of two ATP-binding proteins
 CC (modC), two transmembrane proteins (modB) and a solute-binding
 CC protein (modA) (Probable).
 CC -!- SUBCELLULAR LOCATION: Inner membrane-associated (By similarity).
 CC -!- SIMILARITY: Belongs to the ABC transporter family. ModC subfamily.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AL627268; CAD05231.1; -
 DR EMBL; AE016841; AAO89721.1; -
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR003439; ABC_transporter.
 DR InterPro; IPR004606; Mop.
 DR InterPro; IPR008995; MOP like.
 DR InterPro; IPR005116; TOBE.
 DR Pfam; PF00005; ABC_tran; 1.
 DR Pfam; PF03459; TOBE; 1.
 DR SMART; SM00382; AAA; 1.
 DR TIGRFAMs; TIGR00638; Mop; 1.
 DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
 DR PROSITE; PS00893; ABC_TRANSPORTER_2; 1.
 KW Hydrolase; Transport; Molybdenum; Membrane; Inner membrane;
 KW ATP-binding; Complete proteome.
 FT NP_BIND 31 38 ATP (By similarity).
 SQ SEQUENCE 352 AA; 39111 MW; 30E56C25FD1D6683 CRC64;

Query Match 100.0%; Score 25; DB 1; Length 352;
 Best Local Similarity 100.0%; Pred. No. 30;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
 Db 219 SSVMH 223
 |||||

RESULT 6
 MODC_SALTY STANDARD; PRT; 352 AA.
 AC Q8ZOR6;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Molybdenum import ATP-binding protein modC (EC 3.6.3.29).
 GN MODC OR STW0783.
 OS Salmonella typhimurium.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 OX NCBI_TaxID=602;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
 RX MEDLINE=21534948; PubMed=11677609;
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
 RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
 RT LT2.";
 RL Nature 413:852-856(2001).
 CC -!- FUNCTION: Part of the ABC transporter complex modABC (TC 3.A.1.8)
 CC involved in molybdenum import. Responsible for energy coupling to
 CC the transport system (By similarity).
 CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + molybdate (Out) = ADP + phosphate
 CC + molybdate (In).
 CC -!- SUBUNIT: The complex is composed of two ATP-binding proteins
 CC (modC), two transmembrane proteins (modB) and a solute-binding
 CC protein (modA) (Probable).
 CC -!- SUBCELLULAR LOCATION: Inner membrane-associated (By similarity).
 CC -!- SIMILARITY: Belongs to the ABC transporter family. ModC subfamily.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AE008732; AAL19721.1; --
 StyGene; SG77777; modc.
 InterPro; IPR003593; AAA_ATPase.
 InterPro; IPR003439; ABC_transporter.
 InterPro; IPR004606; MOP.
 InterPro; IPR008995; MOP_like.
 InterPro; IPR005116; TOBE.
 Pfam; PF00005; ABC_tran; 1.
 Pfam; PF03459; TOBE; 1.
 SMART; SM00382; AAA; 1.
 TIGRFAMs; TIGR00638; MOP; 1.
 PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
 PROSITE; PS00893; ABC_TRANSPORTER_2; 1.
 Hydrolase; Transport; Molybdenum; Membrane; Inner membrane;
 ATP-binding; Complete proteome.
 NP_BIND 31 38 ATP (By similarity).
 FT BIND 31 38
 SQ SEQUENCE 352 AA; 39055 MW; 65137E645711AAB5 CRC64;

Query Match 100.0%; Score 25; DB 1; Length 352;
 Best Local Similarity 100.0%; Pred. No. 30;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSVMH 5
 |||||
 Db 219 SSVMH 223

RESULT 7
 ID HM21 DROME STANDARD; PRT; 353 AA.
 AC P09089; Q9VI45;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Zerknullt protein 1 (ZEN-1).
 GN ZEN OR ZEN1 OR Z1 OR CG1046.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86112803; PubMed=2892759;
 RA Rushlow C., Doyle H., Hoey T., Levine M.;
 RT "Molecular characterization of the zerknullt region of the
 Antennapedia gene complex in Drosophila."
 RL Genes Dev. 1:1268-1279(1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RA Celniker S.E., Pfeiffer B., Knafels J., Martin C.H., Mayeda C.A.,
 RA Palazzolo M.J.;
 RT "Complete sequence of the Antennapedia complex of Drosophila."
 RL Submitted (JAN-1999) to the EMBL/GenBank/DDBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RX MEDLINE=20196066; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Frankoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,

RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahike C., Davenport L.B., Davies P.,
 RA De Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Galbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).

CC -!- FUNCTION: Required for the differentiation of the dorsal-ventral
 pattern, and does not appear to be involved in the process of
 segmentation.

CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- SIMILARITY: Contains 1 homeobox domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

EMBL; X68347; CAA48416.1; --
 EMBL; AE001572; AAD19800.1; --
 EMBL; AE003674; AAF54087.1; --
 EMBL; AE003674; AAF54087.1; --
 PIR; A43697; A43697.
 PDB; 1KZ2; 19-JUN-02.
 DR TRANSFAC; T00917; --
 DR FlyBase; FBgn004053; zen.
 DR InterPro; IPR001356; Homeobox.
 DR InterPro; IPR000047; HTH_lambdarepressor.
 DR Pfam; PF00046; homeobox_1.
 DR PRINTS; PR00024; HOMEBOX.
 DR PRINTS; PR00031; HTHREPRESSR.
 DR ProDom; PD000010; Homeobox; 1.
 DR SMART; SM00389; HOX; 1.
 DR PROSITE; PS00027; HOMEBOX_1; 1.
 DR PROSITE; PS00071; HOMEBOX_2; 1.
 KW Transcription regulation; Homeobox; DNA-binding; Nuclear protein;
 KW Developmental protein; 3D-structure.
 FT DNA_BIND 90 149 HOMEBOX.
 SQ SEQUENCE 353 AA; 39302 MW; 1FA64031C160CE2B CRC64;

Query Match 100.0%; Score 25; DB 1; Length 353;
 Best Local Similarity 100.0%; Pred. No. 30;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSVMH 5
 |||||
 Db 2 SSVMH 6

RESULT 8

```
HMZ1_DROSU
ID_HM21_DROSU STANDARD; PRT; 373 AA.
AC Q24648;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Zerknuell protein.
GN ZEN.
OS Drosophila subobscura (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7241;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H271;
RX MEDLINE=96088375; PubMed=7587584;
RA Terol J., Perez-Alonso M., de Prutos R.;
RT "Molecular characterization of the zerknuell region of the
RT Antennapedia complex of D. subobscura.";
RL Chromosoma 103:613-624(1995).
CC -!- FUNCTION: Required for the differentiation of the dorsal-ventral
CC pattern, and does not appear to be involved in the process of
CC segmentation (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- SIMILARITY: Contains 1 homeobox domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X78058; CAA54976.1; -.
CC HSSP; P14653; 1B72.
CC FlyBase; FBgn0015180; Dsub|zen.
CC InterPro; IPR001356; Homeobox.
CC InterPro; IPR000047; HTH_lambdarepressor.
CC Pfam; PF00046; homeobox; 1.
CC PRINTS; PR00024; Homeobox; 1.
CC PRODOM; PD00031; HTHREPRESSR.
CC PRODOM; PD00010; Homeobox; 1.
CC SMART; SM00389; HOX; 1.
CC PROSITE; PS00027; HOMEBOX_1; 1.
CC PROSITE; PS0071; HOMEBOX_2; 1.
CC Transcription regulation; Homeobox; DNA-binding; Nuclear protein;
KW Developmental protein.
FT DNA_BIND 88 147 HOMEBOX
SQ SEQUENCE 373 AA; 40602 MW; 7D647EA2241045F2 CRC64;

Query Match 100.0%; Score 25; DB 1; Length 373;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
Db |||||
2 SSVMH 6

RESULT 9
ID_CYB_CRATY STANDARD; PRT; 379 AA.
AC Q8WDU1; Q8W7F3;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cytochrome b.
GN MTCYB OR COB OR CVTB.
OS Cratogeomys tylosinus (Naked-nosed pocket gopher).
OG Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OG Mitochondrion.
OX NCBI_TaxID=27617;
RN [1]

Mammalia; Euthera; Rodentia; Sciurognathi; Geomyidae; Cratogeomys.
NCBI_TaxID=13463;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Isolate G8, Isolate G30, and Isolate G31;
RX MEDLINE=21655508; PubMed=11796037;
RA Demastes J.W., Spradling T.A., Hafner M.S., Hafner D.J., Reed D.L.;
RT "Systematics and phylogeography of pocket gophers in the genera
RT Cratogeomys and Pappogeomys.";
RL Mol. Phylogenet. Evol. 22:144-154(2002).
CC -!- FUNCTION: Component of the ubiquinol-cytochrome c reductase
CC complex (complex III or cytochrome b-c1 complex), which is a
CC respiratory chain that generates an electrochemical potential
CC coupled to ATP synthesis (By similarity).
CC -!- COFACTOR: Binds two heme groups non-covalently. Heme 1 (or BL or
CC b562) is low-potential and absorbs at about 562, and heme 2 (or BH
CC or b566) is high-potential and absorbs at about 566 (By
CC similarity).
CC -!- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
CC cytochrome c1 and the Rieske protein (By similarity).
CC -!- SIMILARITY: Belongs to the cytochrome b family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF302162; AAL69575.1; -.
CC EMBL; AF302182; AAL69595.1; -.
CC EMBL; AF302183; AAL69596.1; -.
CC InterPro; IPR005798; Cytochrome b C.
CC InterPro; IPR005797; Cytochrome b N.
CC Pfam; PF00032; cytochrome b C; 1.
CC Pfam; PF00033; cytochrome b N; 1.
CC PROSITE; PS00192; CYTOCHROME B HEME; 1.
CC PROSITE; PS00193; CYTOCHROME B_QO; 1.
KW Electron transport; Mitochondrion; Respiratory chain; Transmembrane;
Heme.
FT METAL 83 83 IRON 1 (HEME B562 AXIAL LIGAND) .
FT METAL 97 97 IRON 2 (HEME B566 AXIAL LIGAND) .
FT METAL 182 182 IRON 1 (HEME B562 AXIAL LIGAND) .
FT METAL 196 196 IRON 2 (HEME B566 AXIAL LIGAND) .
FT VARIANT 266 266 P -> S (in isolate G30) .
SQ SEQUENCE 379 AA; 42919 MW; F3EDF1CAFABFAD3 CRC64;

Query Match 100.0%; Score 25; DB 1; Length 379;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
Db |||||
64 SSVMH 68

RESULT 10
ID_CYB_MESPE STANDARD; PRT; 379 AA.
AC Q35066;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cytochrome b.
GN MTCYB OR COB OR CVTB.
OS Mesopodion peruvianus (Peruvian beaked whale).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Cetartiodactyla; Cetacea; Odontoceti; Ziphiidae;
OC Mesopodion.
OX NCBI_TaxID=27617;
RN [1]
```

```

RP SEQUENCE FROM N.A.
RX MEDLINE=20481910; PubMed=11027333;
RA Cassens I., Vicario S., Waddell V.G., Balchowsky H., Van Belle D.,
RA Ding W., Fan C., Mohan L., Simoes-Lopes P.C., Baetida R., Meyer A.,
RA Stanhope M.J., Milinkovitch M.C.;
RT "Independent adaptation to riverine habitats allowed survival of
RT ancient cetacean lineages.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:11343-11347(2000).
RN [2]
RP SEQUENCE OF 1-134 FROM N.A.
RX MEDLINE=95115542; PubMed=7755710;
RA Milinkovitch M.C., Meyer A., Powell J.R.;
RT "Phylogeny of all major groups of cetaceans based on DNA sequences
RT from three mitochondrial genes.";
RL Mol. Biol. Evol. 11:939-948(1994).
CC -!- FUNCTION: Component of the ubiquinol-cytochrome c reductase
CC complex (complex III or cytochrome b-c1 complex), which is a
CC respiratory chain that generates an electrochemical potential
CC coupled to ATP synthesis (By similarity).
CC -!- COFACTOR: Binds two heme groups non-covalently. Heme 1 (or BL or
CC b562) is low-potential and absorbs at about 562, and heme 2 (or BH
CC or b566) is high-potential and absorbs at about 566 (By
CC similarity).
CC -!- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
CC cytochrome c1 and the Rieske protein (By similarity).
CC -!- SIMILARITY: Belongs to the cytochrome b family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF304074; AAC48451.2; -.
DR HSSP; P00157; IBE3.
DR InterPro; IPR005798; Cytb_b6_C.
DR InterPro; IPR005797; Cytb_b6_N.
DR Pfam; PF00032; cytochrome_b_C; 1.
DR Pfam; PF00033; cytochrome_b_N; 1.
DR PROSITE; PS00192; CYTOCHROME_B_HEME; 1.
DR PROSITE; PS00193; CYTOCHROME_B_QO; 1.
DR Electron transport; Mitochondrion; Respiratory chain; Transmembrane;
KW Heme.
FT METAL 83 83 IRON 1 (HEME B562 AXIAL LIGAND).
FT METAL 97 97 IRON 2 (HEME B566 AXIAL LIGAND).
FT METAL 182 182 IRON 1 (HEME B562 AXIAL LIGAND).
FT METAL 196 196 IRON 2 (HEME B566 AXIAL LIGAND).
SQ SEQUENCE 379 AA; 43066 MW; 1BB8AC9AAE5D2050 CRC64;

Query Match 100.0%; Score 25; DB 1; Length 379;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
Db 64 SSVMH 68

RESULT 11
CYB_PERTA
ID CYB_PERTA STANDARD; PRT; 380 AA.
AC Q94T21.
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cytochrome b.
GN MTCYB OR COB OR CYTB.
OS Percopsis transmontana (Sand roller).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Neoteleostei;

```

```

OC Acanthomorpha; Paracanthopterygii; Percopsiformes; Percopsoids;
OC Percopsidae; Percopsis.
OX NCBI_TaxID=143327;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21519020; PubMed=11606696;
RA Miya M., Kawaguchi A., Nishida M.;
RT "Mitogenomic exploration of higher teleostean phylogenies: a case
RT study for moderate-scale evolutionary genomics with 38 newly
RT determined complete mitochondrial DNA sequences.";
RL Mol. Biol. Evol. 18:1993-2009(2001).
CC -!- FUNCTION: Component of the ubiquinol-cytochrome c reductase
CC complex (complex III or cytochrome b-c1 complex), which is a
CC respiratory chain that generates an electrochemical potential
CC coupled to ATP synthesis (By similarity).
CC -!- COFACTOR: Binds two heme groups non-covalently. Heme 1 (or BL or
CC b562) is low-potential and absorbs at about 562, and heme 2 (or BH
CC or b566) is high-potential and absorbs at about 566 (By
CC similarity).
CC -!- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
CC cytochrome c1 and the Rieske protein (By similarity).
CC -!- SIMILARITY: Belongs to the cytochrome b family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AP002928; BAB70156.1; -.
DR InterPro; IPR005798; Cytb_b6_C.
DR InterPro; IPR005797; Cytb_b6_N.
DR Pfam; PF00032; cytochrome_b_C; 1.
DR Pfam; PF00033; cytochrome_b_N; 1.
DR PROSITE; PS00192; CYTOCHROME_B_HEME; 1.
DR PROSITE; PS00193; CYTOCHROME_B_QO; 1.
DR Electron transport; Mitochondrion; Respiratory chain; Transmembrane;
KW Heme.
FT METAL 83 83 IRON 1 (HEME B562 AXIAL LIGAND).
FT METAL 97 97 IRON 2 (HEME B566 AXIAL LIGAND).
FT METAL 182 182 IRON 1 (HEME B562 AXIAL LIGAND).
FT METAL 196 196 IRON 2 (HEME B566 AXIAL LIGAND).
SQ SEQUENCE 380 AA; 42233 MW; 3CE8BCBA74271968 CRC64;

Query Match 100.0%; Score 25; DB 1; Length 380;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
Db 64 SSVMH 68

RESULT 12
CYB_PUNCO
ID CYB_PUNCO STANDARD; PRT; 380 AA.
AC Q9B2C4.
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cytochrome b.
GN MTCYB OR COB OR CYTB.
OS Puntius conchonius (Rosy barb).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteiophysi; Cypriniformes;
OC Cyprinidae; Puntius.
OX NCBI_TaxID=27708;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21097364; PubMed=11161760;

```

RA Machordom A., Doadrio I.;
 RT "Evidence of a cenozoic Betic-Kabilian Connection based on freshwater
 RL fish phylogeography (luccobius, Cyprinidae).";
 CC Mol. Phylogenet. Evol. 18:252-263(2001).
 CC -!- FUNCTION: Component of the ubiquinol-cytochrome c reductase
 CC complex (complex III or cytochrome b-c1 complex), which is a
 CC respiratory chain that generates an electrochemical potential
 CC coupled to ATP synthesis (By similarity).
 CC -!- COFACTOR: Binds two heme groups non-covalently. Heme 1 (or BL or
 CC b562) is low-potential and absorbs at about 562, and heme 2 (or BH
 CC or b566) is high-potential and absorbs at about 566 (By
 CC similarity).
 CC -!- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
 CC cytochrome c1 and the Rieske protein (By similarity).
 CC -!- SIMILARITY: Belongs to the cytochrome b family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AY004751; AAK12998.1; -;
 DR InterPro; IPR005798; Cytb_b6_C.
 DR InterPro; IPR005797; Cytb_b6_N.
 DR Pfam; PF00032; cytochrome_b_c1; 1.
 DR Pfam; PF00033; cytochrome_b_N; 1.
 DR PROSITE; PS00192; CYTOCHROME_B_HEME; 1.
 DR PROSITE; PS00193; CYTOCHROME_B_OO; 1.
 KW Electron transport; Mitochondrion; Respiratory chain; Transmembrane;
 KW Heme.
 FT METAL 83 83 IRON 1 (HEME B562 AXIAL LIGAND).
 FT METAL 97 97 IRON 2 (HEME B566 AXIAL LIGAND).
 FT METAL 182 182 IRON 1 (HEME B562 AXIAL LIGAND).
 FT METAL 196 196 IRON 2 (HEME B566 AXIAL LIGAND).
 FT METAL 196 196 IRON 2 (HEME B566 AXIAL LIGAND).
 SQ SEQUENCE 380 AA; 42732 MW; 7AFA02ECC599306A CRC64;
 Query Match 100.0%; Score 25; DB 1; Length 380;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SSVMH 5
 DB 64 SSVMH 68
 RESULT 13
 CYB_STRPU STANDARD; PRT; 380 AA.
 AC F15547;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Cytochrome b.
 GN MRCYB OR COB OR CYTB.
 OS Strongylocentrotus purpuratus (Purple sea urchin).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
 OC Echinoidea; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae;
 OC Strongylocentrotus.
 OC NCBI_TaxID=7669;
 RN [1]
 RX SEQUENCE FROM N.A.
 RP MEDLINE=89011951; PubMed=3172215;
 RA Jacobs H.T., Elliott D.J., Math V.B., Farquharson A.;
 RT "Nucleotide sequence and gene organization of sea urchin
 RL mitochondrial DNA."; J. Mol. Biol. 202:185-217(1988).
 CC -!- FUNCTION: Component of the ubiquinol-cytochrome c reductase
 CC complex (complex III or cytochrome b-c1 complex), which is a
 CC respiratory chain that generates an electrochemical potential

CC coupled to ATP synthesis (By similarity).
 CC -!- COFACTOR: Binds two heme groups non-covalently. Heme 1 (or BL or
 CC b562) is low-potential and absorbs at about 562, and heme 2 (or BH
 CC or b566) is high-potential and absorbs at about 566 (By
 CC similarity).
 CC -!- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
 CC cytochrome c1 and the Rieske protein (By similarity).
 CC -!- SIMILARITY: Belongs to the cytochrome b family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X12631; CAA31162.1; ALT_INIT.
 DR PIR; S01511; S01511.
 DR InterPro; IPR005798; Cytb_b6_C.
 DR InterPro; IPR005797; Cytb_b6_N.
 DR Pfam; PF00032; cytochrome_b_c1; 1.
 DR Pfam; PF00033; cytochrome_b_N; 1.
 DR PROSITE; PS00192; CYTOCHROME_B_HEME; 1.
 DR PROSITE; PS00193; CYTOCHROME_B_OO; 1.
 KW Electron transport; Mitochondrion; Respiratory chain; Transmembrane;
 KW Heme.
 FT METAL 84 84 IRON 1 (HEME B562 AXIAL LIGAND).
 FT METAL 98 98 IRON 2 (HEME B566 AXIAL LIGAND).
 FT METAL 183 183 IRON 1 (HEME B562 AXIAL LIGAND).
 FT METAL 197 197 IRON 2 (HEME B566 AXIAL LIGAND).
 SQ SEQUENCE 380 AA; 42733 MW; B151A26B55A214F1 CRC64;
 Query Match 100.0%; Score 25; DB 1; Length 380;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SSVMH 5
 DB 65 SSVMH 69
 RESULT 14
 NMT_DROME STANDARD; PRT; 387 AA.
 ID NMT_DROME
 AC O61613;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Glycylpeptide N-tetradecanoyltransferase (EC 2.3.1.97) (Peptide N-
 DE myristoyltransferase) (Myristoyl-CoA:protein N-myristoyltransferase)
 DE (NMT) (dNMT).
 GN NMT.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OC NCBI_TaxID=7227;
 RN [1]
 RX SEQUENCE FROM N.A.
 RP MEDLINE=97196956; PubMed=9044045;
 RA Ntwsa M., Egerton M., Gay N.J.;
 RT "Sequence and expression of Drosophila myristoyl-CoA: protein N-
 RT myristoyl transferase: evidence for proteolytic processing and
 RT membrane localisation."; J. Cell Sci. 110:149-156(1997).
 CC -!- FUNCTION: Adds a myristoyl group to the N-terminal glycine residue
 CC of certain cellular proteins.
 CC -!- CATALYTIC ACTIVITY: Tetradecanoyl-CoA + glycyl-peptide = CoA + N-
 CC tetradecanoylglycyl-peptide.
 CC -!- SUBCELLULAR LOCATION: Membrane-bound.
 CC -!- SIMILARITY: Belongs to the NMT family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

CC -----
CC DR HMBP; AF053725; AAC08578.1; --
CC DR HSSP; P30418; INMT.
CC DR FlyBase; Fgn020392; Nmt.
CC DR InterPro; IPR000903; Nmt.
CC DR Pfam; PF01233; NMT; 1.
CC DR Pfam; PF02799; NMT; 1.
CC DR PROSITE; PS00975; NMT; 1.
CC DR PROSITE; PS00976; NMT; 2.
CC DR Transferrase; Acyltransferase; Membrane.
KW SEQUENCE 387 AA; 45121 MW; 0BF518F24CECEC0C CRC64;
SQ

Query Match 100.0%; Score 25; DB 1; Length 387;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
Db 300 SSVMH 304

RESULT 15

CVB LAMFL
ID CVB LAMFL STANDARD; PRT; 396 AA.
AC Q9TA00;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cytochrome b.
GN MTCYB OR COB OR CVTB.
OS Lampetra fluviatilis (River lamprey).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Lampetra.
OX NCBI_TaxID=7748;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20206853; PubMed=10742044;
RA Delarbre C., Escriva H., Gallut C., Barriol V., Kourilsky P.,
RA Janvier P., Laudet V., Gachet G.;
RT "The complete nucleotide sequence of the mitochondrial DNA of the
RT agnathan Lampetra fluviatilis: bearings on the phylogeny of
RT cyclostomes.";
RL Mol. Biol. Evol. 17:519-529(2000).
CC -!- FUNCTION: Component of the ubiquinol-cytochrome c reductase
CC complex (complex III or cytochrome b-c1 complex), which is a
CC respiratory chain that generates an electrochemical potential
CC coupled to ATP synthesis (By similarity).
CC COFACTOR: Binds two heme groups non-covalently. Heme 1 (or BL or
CC b562) is low-potential and absorbs at about 562, and heme 2 (or BH
CC or b566) is high-potential and absorbs at about 566 (By
CC similarity).
CC -!- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
CC cytochrome c1 and the Rieske protein (By similarity).
CC -!- SIMILARITY: Belongs to the cytochrome b family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

CC -----
CC DR EMBL; Y18683; CAB62226.1; --
CC DR InterPro; IPR005798; Cytb_b6_c.

DR InterPro; IPR005797; Cytb_b6_N.
DR Pfam; PF00032; cytochrome_b_c1.
DR Pfam; PF00033; cytochrome_b_N; 1.
DR PROSITE; PS00192; CYTOCHROME B_HEME; 1.
DR PROSITE; PS00193; CYTOCHROME B_QO; 1.
KW Electron transport; Mitochondrion; Respiratory chain; Transmembrane;
KW Heme.
FT METAL 87 87 IRON 1 (HEME B562 AXIAL LIGAND).
FT METAL 101 101 IRON 2 (HEME B566 AXIAL LIGAND).
FT METAL 186 186 IRON 1 (HEME B562 AXIAL LIGAND).
FT METAL 200 200 IRON 2 (HEME B566 AXIAL LIGAND).
SQ SEQUENCE 396 AA; 44541 MW; B2B69A5CB1C3304C CRC64;

Query Match 100.0%; Score 25; DB 1; Length 396;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
Db 68 SSVMH 72

RESULT 16

CVB PETMA STANDARD; PRT; 396 AA.
AC Q3534;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cytochrome b.
GN MTCYB OR COB OR CVTB.
OS Petromyzon marinus (Sea lamprey).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Petromyzon.
OX NCBI_TaxID=7757;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95229067; PubMed=7713438;
RA Lee W.J., Kocher T.D.;
RT "Complete sequence of a sea lamprey (Petromyzon marinus)
RT mitochondrial genome: early establishment of the vertebrate genome
RT organization.";
RL Genetics 139:873-887(1995).
CC -!- FUNCTION: Component of the ubiquinol-cytochrome c reductase
CC complex (complex III or cytochrome b-c1 complex), which is a
CC respiratory chain that generates an electrochemical potential
CC coupled to ATP synthesis (By similarity).
CC COFACTOR: Binds two heme groups non-covalently. Heme 1 (or BL or
CC b562) is low-potential and absorbs at about 562, and heme 2 (or BH
CC or b566) is high-potential and absorbs at about 566 (By
CC similarity).
CC -!- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
CC cytochrome c1 and the Rieske protein (By similarity).
CC -!- SIMILARITY: Belongs to the cytochrome b family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

CC -----
CC DR EMBL; U11880; AAB08737.1; --

DR PIR; S54999; S54999.
DR InterPro; IPR005798; Cytb_b6_c.
DR InterPro; IPR005797; Cytb_b6_N.
DR Pfam; PF00032; cytochrome_b_c1.
DR Pfam; PF00033; cytochrome_b_N; 1.
DR PROSITE; PS00192; CYTOCHROME B_HEME; 1.
DR PROSITE; PS00193; CYTOCHROME B_QO; 1.
KW Electron transport; Mitochondrion; Respiratory chain; Transmembrane;
KW

KW Heme.
FT METAL 87 87 IRON 1 (HEME B562 AXIAL LIGAND).
FT METAL 101 101 IRON 2 (HEME B566 AXIAL LIGAND).
FT METAL 186 186 IRON 1 (HEME B562 AXIAL LIGAND).
FT METAL 200 200 IRON 2 (HEME B566 AXIAL LIGAND).
SQ SEQUENCE 396 AA; 44551 MW; A7983DA248A0767D CRC64;

Query Match 100.0%; Score 25; DB 1; Length 396;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
DB 68 SSVMH 72

RESULT 17

RHLA_PANTR
ID RHLA_PANTR STANDARD; PRT; 416 AA.
AC Q28813;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE RH-like protein 1A (Rhesus-like protein 1A).
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RX MEDLINE=95085595; PubMed=7993375;
RA Salvignol I., Blancher A., Calvas P., Clayton J., Socha W.W.,
RA Colin Y., Ruffie J.;
RT "Molecular genetics of chimpanzee Rh-related genes: their
RT relationship with the R-C-E-F blood group system, the chimpanzee
RT counterpart of the human rhesus system.";
RL Biochem. Genet. 32:201-221(1994).
CC -!- FUNCTION: MAY BE PART OF AN OLIGOMERIC COMPLEX WHICH IS LIKELY TO
CC HAVE A TRANSPORT OR CHANNEL FUNCTION IN THE ERYTHROCYTE MEMBRANE.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: Belongs to the Rh family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

EMBL; L37049; AAA65623.1; -.
DR PIR; I37004; I37004.
DR InterPro; IPR001905; Ammonium transp.
DR InterPro; IPR002229; RhesusRHD.
DR Pfam; PF00909; Ammonium transp; 1.
DR PRINTS; PR00342; RHESUSRHD.
KW Erythrocyte; Transmembrane.
FT INIT MET 0 0 BY SIMILARITY.
FT TRANSMEM 11 31 POTENTIAL.
FT TRANSMEM 43 63 POTENTIAL.
FT TRANSMEM 76 96 POTENTIAL.
FT TRANSMEM 124 144 POTENTIAL.
FT TRANSMEM 171 191 POTENTIAL.
FT TRANSMEM 202 222 POTENTIAL.
FT TRANSMEM 237 257 POTENTIAL.
FT TRANSMEM 264 284 POTENTIAL.
FT TRANSMEM 286 306 POTENTIAL.
FT TRANSMEM 330 350 POTENTIAL.
FT TRANSMEM 357 377 POTENTIAL.
SQ SEQUENCE 416 AA; 45461 MW; B0B566734DB5E14D CRC64;

Query Match 100.0%; Score 25; DB 1; Length 416;

Best Local Similarity 100.0%; Pred. No. 36;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SSVMH 5
DB 324 SSVMH 328

RESULT 18

RHLC_GORGO
ID RHLC_GORGO STANDARD; PRT; 416 AA.
AC Q28426;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE RH-like protein 1C (Rhesus-like protein 1C).
OS Gorilla gorilla gorilla (Lowland gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
OX NCBI_TaxID=9595;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RX MEDLINE=95085595; PubMed=7993375;
RA Salvignol I., Blancher A., Calvas P., Clayton J., Socha W.W.,
RA Colin Y., Ruffie J.;
RT "Molecular genetics of chimpanzee Rh-related genes: their
RT relationship with the R-C-E-F blood group system, the chimpanzee
RT counterpart of the human rhesus system.";
RL Biochem. Genet. 32:201-221(1994).
CC -!- FUNCTION: MAY BE PART OF AN OLIGOMERIC COMPLEX WHICH IS LIKELY TO
CC HAVE A TRANSPORT OR CHANNEL FUNCTION IN THE ERYTHROCYTE MEMBRANE.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: Belongs to the Rh family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

EMBL; L37052; AAA65626.1; -.
DR PIR; I37075; I37075.
DR InterPro; IPR001905; Ammonium transp.
DR InterPro; IPR002229; RhesusRHD.
DR Pfam; PF00909; Ammonium transp; 1.
DR PRINTS; PR00342; RHESUSRHD.
KW Erythrocyte; Transmembrane.
FT INIT MET 0 0 BY SIMILARITY.
FT TRANSMEM 11 31 POTENTIAL.
FT TRANSMEM 43 63 POTENTIAL.
FT TRANSMEM 76 96 POTENTIAL.
FT TRANSMEM 124 144 POTENTIAL.
FT TRANSMEM 171 191 POTENTIAL.
FT TRANSMEM 202 222 POTENTIAL.
FT TRANSMEM 237 257 POTENTIAL.
FT TRANSMEM 264 284 POTENTIAL.
FT TRANSMEM 286 306 POTENTIAL.
FT TRANSMEM 330 350 POTENTIAL.
FT TRANSMEM 357 377 POTENTIAL.
SQ SEQUENCE 416 AA; 45129 MW; 0F3B623F9908E087 CRC64;

Query Match 100.0%; Score 25; DB 1; Length 416;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
DB 324 SSVMH 328

```
RESULT 19
RHLD_GORGO STANDARD; PRT; 416 AA.
AC Q28427;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 01-NOV-1997 (Rel. 35, Last annotation update)
DE Rhesus-like protein.
OS Gorilla gorilla gorilla (Lowland gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
OX NCBI_TaxID=9595;
RN [1]
SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RA MEDLINE=95085595; PubMed=7993375;
RA Salvignol I., Blancher A., Calvas P., Clayton J., Socha W.W.,
RA Colin Y., Ruffie J.;
RT "Molecular genetics of chimpanzee Rh-related genes: their
RT relationship with the R-C-E-F blood group system, the chimpanzee
RT counterpart of the human rhesus system.";
RL Biochem. Genet. 32:201-221(1994).
CC -!- FUNCTION: MAY BE PART OF AN OLIGOMERIC COMPLEX WHICH IS LIKELY TO
CC HAVE A TRANSPORT OR CHANNEL FUNCTION IN THE ERYTHROCYTE MEMBRANE.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: Belongs to the Rh family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L37048; AAA65622.1; -
CC PIR; I37003; I37003.
CC InterPro; IPR001905; Ammonium transp.
CC InterPro; IPR002229; RhesusRHD.
CC Pfam; PF00909; Ammonium transp; 1.
CC PRINTS; PR00342; RHEUSURHD.
KW Erythrocyte; Transmembrane.
FT INIT MET 0 0 BY SIMILARITY.
FT TRANSMEM 11 31 POTENTIAL.
FT TRANSMEM 43 63 POTENTIAL.
FT TRANSMEM 76 96 POTENTIAL.
FT TRANSMEM 124 144 POTENTIAL.
FT TRANSMEM 171 191 POTENTIAL.
FT TRANSMEM 202 222 POTENTIAL.
FT TRANSMEM 237 257 POTENTIAL.
FT TRANSMEM 264 284 POTENTIAL.
FT TRANSMEM 286 306 POTENTIAL.
FT TRANSMEM 330 350 POTENTIAL.
FT TRANSMEM 357 377 POTENTIAL.
SQ SEQUENCE 416 AA; 44838 MW; 398877F2B7BEE802 CRC64;

Query Match 100.0%; Score 25; DB 1; Length 416;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
DB 324 SSVMH 328
|||||

RESULT 20
RHLD_PANTR STANDARD; PRT; 416 AA.
AC Q28612;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 01-NOV-1997 (Rel. 35, Last annotation update)
DE RH-like protein IIF (Rhesus-like protein IIF).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A. (ISOFORM X1).
RC TISSUE=Thymus;

Query Match 100.0%; Score 25; DB 1; Length 416;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
DB 324 SSVMH 328
|||||

RESULT 21
NFC3_MOUSE STANDARD; PRT; 1075 AA.
AC P97305; Q60896;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Nuclear factor of activated T-cells, cytoplasmic 3 (T cell
DE transcription factor NFAT4) (NF-ATc3) (NF-AT4) (NFATX).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A. (ISOFORM X1).
RC TISSUE=Thymus;
```

RX MEDLINE=95378239; PubMed=7650004;
 RA Ho S.N., Thomas D.J., Timmerman L.A., Li X., Francke U.,
 RA Crabtree G.R.;
 RT "NFATc3, a lymphoid-specific NFATc family member that is
 RT calcium-regulated and exhibits distinct DNA binding specificity.";
 RL J. Biol. Chem. 270:19898-19907(1995).
 RN [2]
 RP SEQUENCE OF 12-1075 FROM N.A. (ISOFORMS X1; X2 AND DELTA-X).
 RC TISSUE=Thymic lymphoma;
 RA Liu J., Koyano-Nakagawa N., Amasaki Y., Saito-Obara F., Ikeuchi T.,
 RA Imai S.-I., Takano T., Arai N., Yokota T., Arai K.-I.;
 RT "Calcineurin-dependent nuclear translocation of a murine transcription
 RT factor NFATx: molecular cloning and functional characterization.";
 RL Mol. Biol. Cell 8:157-170(1997).
 RN [3]
 RP REVIEW.
 RX MEDLINE=99189746; PubMed=10089876;
 RA Crabtree G.R.;
 RT "Generic signals and specific outcomes: signaling through Ca2+,
 RT calcineurin, and NF-AT.";
 RL Cell 96:611-614(1999).
 CC -!- FUNCTION: Plays a role in the inducible expression of cytokine
 CC genes in T cells, especially in the induction of the IL-2 (By
 CC similarity).
 CC -!- SUBUNIT: Member of the multicomponent NFATC transcription complex
 CC that consists of at least two components, a pre-existing
 CC cytoplasmic component NFATC2 and an inducible nuclear component
 CC NFATC1. Other members such as NFAT, NFATC3 or members of the
 CC activating protein-1 family, NFAT, NFAT4 and Cbp/p300 can also bind
 CC the complex. NFATC proteins bind to DNA as monomers.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic for the phosphorylated form and
 CC nuclear after activation that is controlled by calcineurin-
 CC mediated dephosphorylation. Rapid nuclear exit of NFATC is thought
 CC to be one mechanism by which cells distinguish between sustained
 CC and transient calcium signals. The subcellular localization of
 CC NFATC play a key role in the gene transcription.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=3;
 CC Name=X1;
 CC IsoId=P97305-1; Sequence=Displayed;
 CC Name=X2;
 CC IsoId=P97305-2; Sequence=VSP_005604;
 CC Name=Delta-X;
 CC IsoId=P97305-3; Sequence=VSP_005603;
 CC -!- TISSUE SPECIFICITY: Expressed in thymus. Weakly expressed in
 CC muscle, spleen and kidney. Also expressed in lymph node.
 CC -!- DOMAIN: Rel Similarity Domain (RSD) allows DNA-binding and
 CC cooperative interactions with AP1 factors (By similarity).
 CC -!- PTM: Phosphorylated by NFATC-kinase; dephosphorylated by
 CC calcineurin (By similarity).
 CC -!- SIMILARITY: Belongs to the Rel/Dorsal family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; D85612; BAAL2833.1; -;
 CC EMBL; U28807; AAA93249.1; -;
 CC TRANSFAC; T01949; -;
 CC MGD; MGI:103296; NFatc3.
 CC GO; GO:0045333; P:cellular respiration; IMP.
 CC GO; GO:0007507; P:heart development; IGI.
 CC GO; GO:0006355; P:regulation of transcription, DNA-dependent; IMP.
 CC InterPro; IPR007110; IG-like.
 CC InterPro; IPR002909; IPT TIG.
 CC InterPro; IPR000451; NF Rel_dor.
 CC InterPro; IPR008366; NFAT.
 CC InterPro; IPR008967; P53-like.

DR Pfam: PF00554; RHD; 1.
 DR Pfam: PF01833; TIG; 1.
 DR PRINTS; PRO1789; NUCTACTORATC.
 DR SMART; SM00429; IPT; 1.
 DR PROSITE; PS01204; REL_1; FALSE_NEG.
 DR PROSITE; PS0254; REL_2; 1.
 KW Transcription regulation; Activator; Nuclear protein; DNA-binding;
 KW Alternative splicing; Repeat; Phosphorylation.
 FT DOMAIN 24 29
 FT POLY-PRO.
 FT DOMAIN 109 114
 FT CALCINEURIN-BINDING.
 FT DOMAIN 207 308
 FT 3 X SP REPEATS.
 FT REPEAT 207 223
 FT REPEAT 236 252
 FT REPEAT 292 308
 FT REPEAT 273 275
 FT DOMAIN 444 451
 FT NUCLEAR LOCALIZATION SIGNAL.
 FT DOMAIN 686 688
 FT NUCLEAR EXPORT SIGNAL.
 FT DOMAIN 1031 1040
 FT Missing (in isoform Delta-X).
 FT VARSPLIC 468 497
 FT /FTid=VSP_005603.
 FT VNEIGRDMSQISVSQATEVMRDTELPASPDLMTSHSAH
 FT -> DOLISDLBHQSGSTKWNHSEFSCPVPFWRI (in
 FT isoform X2).
 FT /FTid=VSP_005604.
 FT LDFKLVFGEDEGAPAPPSPSRPA -> PRRVLFVSVAQLPS
 FT ETRPGPSDL (IN REF. 2)
 FT TTPLC -> NSSLG (IN REF. 2).
 FT P -> LQ (IN REF. 2).
 FT SH -> GY (IN REF. 2).
 FT CEIP -> GDIS (IN REF. 2).
 FT F -> I (IN REF. 2).
 FT QQEL -> HQCT (IN REF. 2).
 FT Q -> H (IN REF. 2).
 FT F -> Y (IN REF. 2).
 FT G -> R (IN REF. 2).
 FT V -> A (IN REF. 2).
 FT REDTLPSPVSLPVPFSAQAQPSSETG -> KENKTOICLQ
 FT FHLCLILPSPEALLRDR (IN REF. 2).
 FT G -> S (IN REF. 2).
 FT H -> Q (IN REF. 2).
 FT SEQUENCE 1075 AA; 115450 MW; 4ED38C9AA6F452BB CRC64;
 Query Match 100.0%; Score 25; DB 1; Length 1075;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SSVHM 5
 Db 783 SSVHM 787
 RESULT 22
 LRP2 RAT
 ID LRP2 RAT STANDARD; PRT; 4660 AA.
 AC P98158;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Low-density lipoprotein receptor-related protein 2 precursor (Megalin)
 DE (Glycoprotein 330) (gp330).
 GN LRP2.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Kidney;
 RX MEDLINE=95024033; PubMed=7937880;
 RA Saito A., Pietromonaco S., Loo A.K.C., Farquhar M.G.;
 RT "Complete cloning and sequencing of rat gp330/megalin, a
 RT distinctive member of the low density lipoprotein receptor gene
 RT family.";

Proc. Natl. Acad. Sci. U.S.A.	91:9725-9729(1994).	FT	TRANSNEM	4426	4446	POTENTIAL.
RL	FUNCTION.	FT	DOMAIN	4426	4446	CYTOPLASMIC (POTENTIAL).
RP	MEDLINE=95386696; PubMed=7544804;	FT	DOMAIN	4447	4660	LDL-RECEPTOR CLASS A 1.
RA	Moestrup S.K., Cui S., Vorum H., Bregengaard C., Bjorn S.E.,	FT	DOMAIN	64	64	LDL-RECEPTOR CLASS A 2.
RA	Norris K., Gliemann J., Christensen E.I.;	FT	DOMAIN	65	105	LDL-RECEPTOR CLASS A 3.
RA	"Evidence that epithelial glycoprotein 330/megalin mediates uptake of	FT	DOMAIN	106	144	LDL-RECEPTOR CLASS A 4.
RT	polybasic drugs.";	FT	DOMAIN	145	181	LDL-RECEPTOR CLASS A 5.
RT	J. Clin. Invest. 96:1404-1413(1995).	FT	DOMAIN	181	219	LDL-RECEPTOR CLASS A 6.
RL	[3]	FT	DOMAIN	220	258	LDL-RECEPTOR CLASS A 7.
RP	TISSUE SPECIFICITY.	FT	DOMAIN	263	307	EGF-LIKE 1.
RP	MEDLINE=94172242; PubMed=7510321;	FT	DOMAIN	308	346	EGF-LIKE 2.
RA	Zheng G., Bachinsky D.R., Stamenkovic I., Strickland D.K., Brown D.,	FT	DOMAIN	347	385	LDL-RECEPTOR CLASS B 1.
RA	Andres G., McCluskey R.T.;	FT	DOMAIN	435	476	LDL-RECEPTOR CLASS B 2.
RT	"Organ distribution in rats of two members of the low-density	FT	DOMAIN	478	519	LDL-RECEPTOR CLASS B 3.
RT	lipoprotein receptor gene family, gp330 and LRP/alpha 2MR, and the	FT	DOMAIN	521	566	LDL-RECEPTOR CLASS B 4.
RT	receptor-associated protein (RAP).";	FT	DOMAIN	568	611	LDL-RECEPTOR CLASS B 5.
RL	J. Histochem. Cytochem. 42:531-542(1994).	FT	DOMAIN	612	652	EGF-LIKE 3.
CC	-!- FUNCTION: Binds plasminogen, extracellular matrix components,	FT	DOMAIN	658	704	LDL-RECEPTOR CLASS B 6.
CC	plasminogen activator-plasminogen activator inhibitor type I	FT	DOMAIN	752	793	LDL-RECEPTOR CLASS B 7.
CC	complex, apolipoprotein E-enriched beta-VLDL, lipoprotein lipase,	FT	DOMAIN	837	879	LDL-RECEPTOR CLASS B 8.
CC	lactoferrin, clusterin and calcium.	FT	DOMAIN	881	923	LDL-RECEPTOR CLASS B 9.
CC	-!- FUNCTION: Receptor-mediated uptake of polybasic drugs such as	FT	DOMAIN	969	1013	EGF-LIKE 4.
CC	aprotinin, aminoglycosides and polymyxin B.	FT	DOMAIN	1023	1061	LDL-RECEPTOR CLASS A 8.
CC	-!- SUBUNIT: Forms a multimeric complex together with a receptor-	FT	DOMAIN	1064	1103	LDL-RECEPTOR CLASS A 9.
CC	associated protein (RAP). Binds to ankryin-repeat family A protein	FT	DOMAIN	1108	1146	LDL-RECEPTOR CLASS A 10.
CC	2 (ANKRA2) (By similarity).	FT	DOMAIN	1148	1186	LDL-RECEPTOR CLASS A 11.
CC	-!- SUBCELLULAR LOCATION: Type I membrane protein. Expressed in	FT	DOMAIN	1186	1225	LDL-RECEPTOR CLASS A 12.
CC	clathrin-coated pits; a soluble form is possibly derived by	FT	DOMAIN	1229	1269	LDL-RECEPTOR CLASS A 13.
CC	cleavage at the cell surface.	FT	DOMAIN	1270	1308	LDL-RECEPTOR CLASS A 14.
CC	-!- TISSUE SPECIFICITY: Epithelial cells of kidney glomerulus and	FT	DOMAIN	1311	1351	LDL-RECEPTOR CLASS A 15.
CC	proximal tubule, lung, epididymis, yolk sac, among others.	FT	DOMAIN	1350	1390	EGF-LIKE 5.
CC	-!- SIMILARITY: Contains 36 LDL-receptor class A domains.	FT	DOMAIN	1391	1430	EGF-LIKE 6.
CC	-!- SIMILARITY: Contains 37 LDL-receptor class B domains.	FT	DOMAIN	1479	1520	LDL-RECEPTOR CLASS B 10.
CC	-!- SIMILARITY: Contains 17 EGF-like domains.	FT	DOMAIN	1522	1563	LDL-RECEPTOR CLASS B 11.
CC	-----	FT	DOMAIN	1567	1609	LDL-RECEPTOR CLASS B 12.
CC	-----	FT	DOMAIN	1611	1654	LDL-RECEPTOR CLASS B 13.
CC	-----	FT	DOMAIN	1656	1696	LDL-RECEPTOR CLASS B 14.
CC	-----	FT	DOMAIN	1701	1742	EGF-LIKE 7.
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	FT	DOMAIN	1791	1832	LDL-RECEPTOR CLASS B 15.
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	FT	DOMAIN	1832	1882	LDL-RECEPTOR CLASS B 16.
CC	the European Bioinformatics Institute. There are no restrictions on its	FT	DOMAIN	1882	1930	LDL-RECEPTOR CLASS B 17.
CC	use by non-profit institutions as long as its content is in no way	FT	DOMAIN	1930	1972	LDL-RECEPTOR CLASS B 18.
CC	modified and this statement is not removed. Usage by and for commercial	FT	DOMAIN	1972	2013	LDL-RECEPTOR CLASS B 19.
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/	FT	DOMAIN	2013	2060	EGF-LIKE 8.
CC	or send an email to license@isb-sib.ch).	FT	DOMAIN	2060	2156	LDL-RECEPTOR CLASS B 20.
CC	-----	FT	DOMAIN	2156	2201	LDL-RECEPTOR CLASS B 21.
CC	-----	FT	DOMAIN	2201	2245	LDL-RECEPTOR CLASS B 22.
CC	-----	FT	DOMAIN	2245	2289	LDL-RECEPTOR CLASS B 23.
CC	-----	FT	DOMAIN	2289	2332	LDL-RECEPTOR CLASS B 24.
CC	-----	FT	DOMAIN	2332	2384	EGF-LIKE 9.

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
DR EMBL; U65377; AAB06740.1; --
DR InterPro; IPR005798; Cytb_b6_C.
DR InterPro; IPR005797; Cytb_b6_N.
DR Pfam; PF00033; cytochrome b_N; 1.
DR PROSITE; PS00192; CYTOCHROME_B_HEME; FALSE NEG.
DR PROSITE; PS00193; CYTOCHROME_B_OO; PARTIAL_
KW Electron transport; Mitochondrion; Respiratory chain; Transmembrane;
KW Heme.
FT NON_TER 1 1
FT METAL 75 75 IRON 1 (HEME B562 AXIAL LIGAND).
FT METAL 89 89 IRON 2 (HEME B566 AXIAL LIGAND).
FT METAL 124 124 IRON 1 (HEME B562 AXIAL LIGAND).
FT NON_TER 128 128
SQ SEQUENCE 128 AA; 14602 MW; 993267807EB97550 CRC64;

Query Match 96.0%; Score 24; DB 1; Length 128;
Best Local Similarity 80.0%; Pred. No. 19;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
||:|
Db 56 SSIMH 60

RESULT 25

CYB_ATRMI STANDARD; PRT; 214 AA.
AC P87421;
DT 15-JUL-1998 (Rel. 36, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cytochrome b (Fragment).
GN MTCYB OR COB OR CYTB.
OS Atractaspis micropholis (Mole viper).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Atractaspididae; Atractaspis.
OX NCBI_TaxID=55260;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=98334559; PubMed=9667984;
RA Vidal N., Lecointre G.;
RT "Weighting and congruence: a case study based on three mitochondrial
genes in pitvipers."
RL Mol. Phylogenet. Evol. 9:366-374(1998).
RN [2]
RP SEQUENCE OF 1-132 FROM N.A.
RA Vidal N., Lecointre G., Vie J.-C., Gasc J.-P.;
RT "Molecular systematics of pitvipers: paraphyly of the Bothrops
complex."
RL C. R. Acad. Sci., III, Sci. Vie 320:95-101(1997).
CC -!- FUNCTION: Component of the ubiquinol-cytochrome c reductase
complex (complex III or cytochrome b-c1 complex), which is a
respiratory chain that generates an electrochemical potential
coupled to ATP synthesis (By similarity).
CC -!- COFACTOR: Binds two heme groups non-covalently. Heme 1 (or BL or
b562) is low-potential and absorbs at about 562, and heme 2 (or BH
or b566) is high-potential and absorbs at about 566 (By
similarity).
CC -!- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
cytochrome c1 and the Rieske protein (By similarity).
CC -!- SIMILARITY: Belongs to the cytochrome b family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
DR EMBL; AF039261; AAC33538.1; --
DR InterPro; IPR005798; Cytb_b6_C.
DR InterPro; IPR005797; Cytb_b6_N.
DR Pfam; PF00033; cytochrome b_N; 1.
DR PROSITE; PS00192; CYTOCHROME_B_HEME; FALSE NEG.
DR PROSITE; PS00193; CYTOCHROME_B_OO; PARTIAL_
KW Electron transport; Mitochondrion; Respiratory chain; Transmembrane;
KW Heme.
FT NON_TER 1 1
FT METAL 81 81 IRON 1 (HEME B562 AXIAL LIGAND).
FT METAL 95 95 IRON 2 (HEME B566 AXIAL LIGAND).
FT NON_TER 214 214
SQ SEQUENCE 214 AA; 24059 MW; 4479973B77768242 CRC64;

Query Match 96.0%; Score 24; DB 1; Length 214;
Best Local Similarity 80.0%; Pred. No. 33;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
||:|
Db 62 SSIMH 66

Search completed: September 24, 2004, 01:55:21
Job time : 8.15094 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 01:47:17 ; Search time 23.9623 Seconds
(without alignments)
65.836 Million cell updates/sec

Title: US-10-088-639A-2 COPY 158 162

Perfect score: 25
Sequence: 1 SSMVH 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters:

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

```
Database : STREMBL_25.*
1:  sp_archea.*
2:  sp_bacteria.*
3:  sp_fungi.*
4:  sp_human.*
5:  sp_invertebrate.*
6:  sp_mammal.*
7:  sp_mhc.*
8:  sp_organelle.*
9:  sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_tvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*
```

Result No.	Score	Query %		Length	DB	ID	Description
		Match					
1	25	100.0	78	2	Q8VM19	Q8VM19 pseudomonas	
2	25	100.0	90	13	Q9PW02	Q9PW02 petromyzon	
3	25	100.0	90	13	Q9Y072	Q9Y072 lampetra pl	
4	25	100.0	93	8	Q35383	Q35383 pantodon bu	
5	25	100.0	100	4	Q9Y347	Q9Y347 homo sapien	
6	25	100.0	102	8	Q9XLK3	Q9XLK3 buglossidii	
7	25	100.0	102	8	Q9B378	Q9B378 dicologloss	
8	25	100.0	110	8	Q9T9U6	Q9T9U6 puntius con	
9	25	100.0	120	8	Q99071	Q99071 pronothobra	
10	25	100.0	120	8	Q99070	Q99070 fundulusoma	
11	25	100.0	128	8	Q9BT72	Q9BT72 lampetra fl	
12	25	100.0	128	8	Q8MIH7	Q8MIH7 mesopiodon	
13	25	100.0	128	8	Q9T3E3	Q9T3E3 entosphenus	
14	25	100.0	128	8	Q9TB73	Q9TB73 lampetra ri	
15	25	100.0	128	8	Q9TB74	Q9TB74 lampetra ay	
16	25	100.0	128	8	Q9TB69	Q9TB69 entosphenus	

SUMMARIES

```

90 24 96.0 225 8 Q9MIH2
91 24 96.0 238 8 Q7Y9Z9
92 24 96.0 238 8 Q7Y9V4
93 24 96.0 242 8 Q85NV5
94 24 96.0 247 3 Q12531
95 24 96.0 254 17 Q97318
96 24 96.0 266 13 Q13116
97 24 96.0 270 5 Q62243
98 24 96.0 340 16 Q98L20
99 24 96.0 353 13 Q9YHY3
100 24 96.0 355 4 Q8N8C0
101 24 96.0 371 8 Q8WA13
102 24 96.0 372 8 Q94UD9
103 24 96.0 420 5 Q81D17
104 24 96.0 469 16 Q81C05
105 24 96.0 471 10 Q9FHK1
106 24 96.0 496 16 Q9CBA3
107 24 96.0 508 11 Q80XN1
108 24 96.0 551 10 Q8H3W7
109 24 96.0 553 16 Q81PG9
110 24 96.0 556 12 Q89180
111 24 96.0 556 12 Q89232
112 24 96.0 590 4 Q96G63
113 24 96.0 626 16 Q8CMP0
114 24 96.0 627 10 Q7Y1L5
115 24 96.0 627 16 Q8G1M1
116 24 96.0 636 4 Q8H9B6
117 24 96.0 646 16 Q8YGA2
118 24 96.0 683 16 Q7UHW4
119 24 96.0 695 4 Q9H062
120 24 96.0 747 4 Q9NVG7
121 24 96.0 747 4 Q96AE5
122 24 96.0 748 11 Q7TMQ7
123 24 96.0 798 3 Q870S0
124 24 96.0 892 12 Q9QN10
125 24 96.0 892 12 Q8V2P5
126 24 96.0 892 12 Q8JF50
127 24 96.0 936 5 Q81FX2
128 24 96.0 951 5 Q20176
129 24 96.0 1209 16 Q8F3P1
130 24 96.0 1475 10 Q9XEP3
131 24 96.0 1489 16 Q8KPC6
132 24 96.0 1553 10 Q8S479
133 24 96.0 34 4 Q9Y4Z8
134 24 96.0 34 16 Q81S09
135 24 96.0 39 16 Q87RL5
136 24 96.0 46 16 Q8EZM6
137 24 96.0 68 10 Q93303
138 24 96.0 68 10 Q93304
139 24 96.0 71 16 Q7VMK3
140 24 96.0 71 16 Q7UTX9
141 24 96.0 78 2 Q9EU09
142 24 96.0 79 16 Q8ZD04
143 24 96.0 86 4 Q8NED2
144 24 96.0 94 2 Q7WUC6
145 24 96.0 97 9 Q8SCL1
146 24 96.0 102 4 Q86TJ8
147 24 96.0 104 5 Q9YIU5
148 24 96.0 104 17 Q9S969
149 24 96.0 113 2 Q53419
150 24 96.0 113 17 Q9YAJ8

```

ALIGNMENTS

```

RESULT 1
Q8VML9 PRELIMINARY; PRT; 78 AA.
ID Q8VML9;
AC Q8VML9; (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)

```

```

DE Hypothetical protein.
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A.
RA Greated A., Lamberton L., Williams P.A., Thomas C.M.;
RT "Complete nucleotide sequence of IncP-9 plasmid pMW0.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ344068; CAC86770.1; -
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 78 AA; 8805 MW; 8762D76ACE5FF30A CRC64;

```

```

Query Match 100.0%; Score 25; DB 2; Length 78;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 SSVMH 5
DB 59 SSVMH 63

```

```

RESULT 2
Q9PW02 PRELIMINARY; PRT; 90 AA.
ID Q9PW02;
AC Q9PW02;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Cytochrome b (Fragment).
GN CYTB OR CYT B.
OS Petromyzon marinus (Sea lamprey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Petromyzon.
OX NCBI_TaxID=7757;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RA Tagliavini J., Tizi R., Conterio F., Mariottini P., Gandolfi G.;
RT "Mitochondrial DNA sequences in three genera of Italian lampreys.";
RL Boll. Zool. 61:331-333(1994).
CC - SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
DR EMBL; X79111; CAA55728.2; -
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016491; P:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005797; Cytb_b6_N.
DR Pfam; PF00033; cytochrome_b_N; 1.
DR PROSITE; PS00192; CYTOCHROME_B_HEME; 1.
KW Electron transport; Heme; Transmembrane; Transport.
FT NON_TER 1
FT NON_TER 90 90
SQ SEQUENCE 90 AA; 10103 MW; AAAED2A254889404 CRC64;

```

```

Query Match 100.0%; Score 25; DB 13; Length 90;
Best Local Similarity 100.0%; Pred. No. 73;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 SSVMH 5
DB 29 SSVMH 33

```

```

RESULT 3
Q9Y072 PRELIMINARY; PRT; 90 AA.
ID Q9Y072
AC Q9Y072;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)

```

DR 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cytochrome b (Fragment).
GN CYTB OR CYT B.
OS Lampetra planeri (Brook lamprey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Lampetra.
OX NCBI_TaxID=7750;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RA Tagliavini J., Tizi R., Conterio F., Mariottini P., Gandolfi G.;
RT "Mitochondrial DNA sequences in three genera of Italian lampreys.";
RL Boll. Zool. 61:331-333(1994).
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
DR EMBL; X79110; CAA55727.2; -.
DR GO; GO:0016021; C-integral to membrane; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005797; Cytb b6 N.
DR Pfam; PF00033; cytochrome b_N; 1.
DR PROSITE; PS00192; CYTOCHROME_B_HEME; 1.
KW Electron transport; Heme; Transmembrane; Transport.
FT NON_TER 1
FT NON_TER 90 90
SQ SEQUENCE 90 AA; 10112 MW; ACOC7153062B63F4 CRC64;

Query Match 100.0%; Score 25; DB 13; Length 90;
Best Local Similarity 100.0%; Pred. No. 73;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
DB 29 SSVMH 33
|||||
RESULT 4
ID Q35383 PRELIMINARY; PRT; 93 AA.
AC Q35383;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cytochrome b (Fragment).
GN CYTB.
OS Pantodon bucholtzi (Butterflyfish).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
OC Osteoglossiformes; Pantodontidae; Pantodon.
OX NCBI_TaxID=8276;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92130804; PubMed=1663569;
RA Nørmark B.B., McCune A.R., Harrison R.G.;
RT "Phylogenetic relationships of neopterygian fishes, inferred from
RT mitochondrial DNA sequences.";
RL Mol. Biol. Evol. 8:819-834(1991).
CC -!- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
CC COUPLED TO ATP SYNTHESIS (BY SIMILARITY).
CC -!- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
CC BOUND TO THE PROTEIN (BY SIMILARITY).
CC -!- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
CC CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
DR EMBL; M64907; AAB01609.1; -.
DR GO; GO:0016021; C-integral to membrane; IEA.
DR GO; GO:0005746; C-mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C-mitochondrion; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.

DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005797; Cytb b6 N.
DR Pfam; PF00033; cytochrome b_N; 1.
KW Electron transport; Heme; Respiratory chain; Transmembrane; Transport;
KW Mitochondrion.
FT NON_TER 1
FT NON_TER 93 93
SQ SEQUENCE 93 AA; 10689 MW; F32P96584CDF43F0 CRC64;

Query Match 100.0%; Score 25; DB 8; Length 93;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
DB 32 SSVMH 36
|||||
RESULT 5
ID Q9Y347 PRELIMINARY; PRT; 100 AA.
AC Q9Y347;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Persson A.E., Lundberg J., Uhlen M.;
RT "EU-IMAGE: Full-insert length sequencing of human cDNA clones.";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF143328; AAD33911.1; -.
DR GenBank; HGNC:15224; OR52K3P.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 100 AA; 11028 MW; 34EAE445D3DBC50A CRC64;

Query Match 100.0%; Score 25; DB 4; Length 100;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
DB 43 SSVMH 47
|||||
RESULT 6
ID Q9XLK3 PRELIMINARY; PRT; 102 AA.
AC Q9XLK3;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cytochrome b (Fragment).
GN CYTB.
OS Buglossidium luteum.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
OC Soleiidae; Soleiidae; Buglossidium.
OX NCBI_TaxID=90071;
RN [1]
RP SEQUENCE FROM N.A.
RA Tinti F., Nadiani M., Tommasini S., Piccinetti C., Stagni A.M.;
RT "Phylogeny and Evolution of Mediterranean Soles (Soleidae,
RT Pleuronectiformes) Based on Mitochondrial DNA Sequences.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE

CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
 CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
 CC COUPLED TO ATP SYNTHESIS (BY SIMILARITY).
 CC -!- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
 CC BOUND TO THE PROTEIN (BY SIMILARITY).
 CC -!- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
 CC CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
 DR EMBL; AF113206; AAD23435.1; -.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO; GO:0005739; C:mitochondrion; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR005797; Cytb b6 N.
 DR Pfam; PF00033; cytochrome b_N; 1.
 DR PROSITE; PS00192; CYTOCHROME B HEME; 1.
 KW Electron transport; Heme; Respiratory chain; Transmembrane; Transport;
 KW Mitochondrion.
 FT NON_TER 1
 FT NON_TER 102
 SQ SEQUENCE 102 AA; 11356 MW; 90F4444E2DD2E4E3 CRC64;
 Query Match 100.0%; Score 25; DB 8; Length 102;
 Best Local Similarity 100.0%; Pred. No. 82;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SSVMH 5
 Db |||||
 32 SSVMH 36

RESULT 7
 Q9B378 PRELIMINARY; PRT; 102 AA.
 ID Q9B378;
 AC Q9B378;
 DT 01-JUN-2001 (TRENBLrel. 17, Created)
 DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
 DE Cytochrome b (Fragment).
 GN CYTB.
 OS Dicloglossa cuneata.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
 OC Soleidae; Soleidae; Dicloglossa.
 OX NCBI_TaxID=153209;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Cunha R., Caetano S., Castilho R.;
 RT "Molecular systematics of Pleuronectiformes based on mitochondrial
 RT DNA";
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
 CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
 CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
 CC COUPLED TO ATP SYNTHESIS (BY SIMILARITY).
 CC -!- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
 CC BOUND TO THE PROTEIN (BY SIMILARITY).
 CC -!- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
 CC CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
 DR EMBL; AF338429; AAK30100.1; -.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO; GO:0005739; C:mitochondrion; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR005797; Cytb b6 N.
 DR InterPro; IPR000991; GATase_1.

DR Pfam; PF00033; cytochrome_b_N; 1.
 DR PROSITE; PS00192; CYTOCHROME B HEME; 1.
 DR PROSITE; PS00442; GATASE TYPE I; 1.
 KW Electron transport; Heme; Respiratory chain; Transmembrane; Transport;
 KW Mitochondrion.
 FT NON_TER 1
 FT NON_TER 102
 SQ SEQUENCE 102 AA; 11318 MW; C89116AD5009B699 CRC64;
 Query Match 100.0%; Score 25; DB 8; Length 102;
 Best Local Similarity 100.0%; Pred. No. 82;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SSVMH 5
 Db |||||
 32 SSVMH 36

RESULT 8
 Q9T9U6 PRELIMINARY; PRT; 110 AA.
 ID Q9T9U6;
 AC Q9T9U6;
 DT 01-MAY-2000 (TRENBLrel. 13, Created)
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
 DE Cytochrome b (Fragment).
 GN CYTB OR CYT B.
 OS Puntius conchionus (Rosy barb).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Osteiophysi; Cypriniformes;
 OC Cyprinidae; Puntius.
 OX NCBI_TaxID=27708;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Gilles A., Lecointre G.;
 RT "A more accurate phylogeny of european cyprinids based on the
 RT mitochondrial control region.";
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
 CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
 CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
 CC COUPLED TO ATP SYNTHESIS (BY SIMILARITY).
 CC -!- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
 CC BOUND TO THE PROTEIN (BY SIMILARITY).
 CC -!- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
 CC CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
 DR EMBL; AJ388455; CAB62077.1; -.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO; GO:0005739; C:mitochondrion; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR005797; Cytb b6 N.
 DR Pfam; PF00033; cytochrome_b_N; 1.
 DR PROSITE; PS00192; CYTOCHROME B HEME; 1.
 KW Electron transport; Heme; Respiratory chain; Transmembrane; Transport;
 KW Mitochondrion.
 FT NON_TER 1
 FT NON_TER 110
 SQ SEQUENCE 110 AA; 12598 MW; FAA7FE8DDA70AC3 CRC64;
 Query Match 100.0%; Score 25; DB 8; Length 110;
 Best Local Similarity 100.0%; Pred. No. 89;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SSVMH 5
 Db |||||
 33 SSVMH 37

```

RESULT 9
O99071 ID Q99071 PRELIMINARY; PRT; 120 AA.
AC O99071;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DE Cytochrome b (Fragment).
GN CYTB.
OS Pronothobranchius kiyawensis.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percormorpha; Atherinomorpha;
OC Cyprinodontiformes; Aplocheilidae; Aplocheilinae; Pronothobranchius.
OX NCBI_TaxID=60306;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Aquarium strain;
RX MEDLINE=99214037; PubMed=10196077;
RA Murphy W.J., Collier G.E.;
RT "Phylogenetic relationships of African killifishes in the genera
RT Aphyosennion and Fundulopanchax inferred from mitochondrial DNA
sequences.";
RT Sequences.
RL Mol. Phylogenet. Evol. 11:351-360(1999).
CC -!- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
CC COUPLED TO ATP SYNTHESIS (BY SIMILARITY).
CC -!- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
CC BOUND TO THE PROTEIN (BY SIMILARITY).
CC -!- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
CC CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
DR EMBL; AF002285; AAD00946.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005797; Cytb_b6_N.
DR Pfam; PF00033; cytochrome_b_N; 1.
DR PROSITE; PS00192; CYTOCHROME_B_HEME; 1.
DR KW Electron transport; Heme; Respiratory chain; Transmembrane; Transport;
KW Mitochondrion.
FT NON_TER 1
FT NON_TER 120
SQ SEQUENCE 120 AA; 13591 MW; 54DIF7D50C1538CE CRC64;

Query Match 100.0%; Score 25; DB 8; Length 120;
Best Local Similarity 100.0%; Pred. No. 96;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
Db 60 SSVMH 64
|||||

RESULT 10
O99070 ID Q99070 PRELIMINARY; PRT; 120 AA.
AC O99070;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DE Cytochrome b (Fragment).
GN CYTB.
OS Fundulosoma thierrii.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percormorpha; Atherinomorpha;
OX NCBI_TaxID=7748;
RN [1]
RP SEQUENCE FROM N.A.
RC Döcker M.F., Youson J.H., Beamish R.J., Devlin R.H.;
RT "Phylogeny of the lamprey genus Lampetra inferred from mitochondrial
RT cytochrome b and ND3 gene sequences.";
RL Can. J. Fish. Aquat. Sci. 56:2340-2349(1999).
CC -!- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
CC COUPLED TO ATP SYNTHESIS (BY SIMILARITY).
CC -!- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
CC BOUND TO THE PROTEIN (BY SIMILARITY).
CC -!- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
CC CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
DR EMBL; AF002285; AAD00946.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005797; Cytb_b6_N.
DR Pfam; PF00033; cytochrome_b_N; 1.
DR PROSITE; PS00192; CYTOCHROME_B_HEME; 1.
DR KW Electron transport; Heme; Respiratory chain; Transmembrane; Transport;
KW Mitochondrion.
FT NON_TER 1
FT NON_TER 120
SQ SEQUENCE 120 AA; 13591 MW; 54DIF7D50C1538CE CRC64;

Query Match 100.0%; Score 25; DB 8; Length 120;
Best Local Similarity 100.0%; Pred. No. 96;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
Db 60 SSVMH 64
|||||

```

```

OC Cyprinodontiformes; Aplocheilidae; Aplocheilinae; Fundulosoma.
OX NCBI_TaxID=60304;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Aquarium strain;
RX MEDLINE=99214037; PubMed=10196077;
RA Murphy W.J., Collier G.E.;
RT "Phylogenetic relationships of African killifishes in the genera
RT Aphyosennion and Fundulopanchax inferred from mitochondrial DNA
sequences.";
RT Sequences.
RL Mol. Phylogenet. Evol. 11:351-360(1999).
CC -!- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
CC COUPLED TO ATP SYNTHESIS (BY SIMILARITY).
CC -!- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
CC BOUND TO THE PROTEIN (BY SIMILARITY).
CC -!- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
CC CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
DR EMBL; AF002284; AAD00945.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005797; Cytb_b6_N.
DR Pfam; PF00033; cytochrome_b_N; 1.
DR PROSITE; PS00192; CYTOCHROME_B_HEME; 1.
DR KW Electron transport; Heme; Respiratory chain; Transmembrane; Transport;
KW Mitochondrion.
FT NON_TER 1
FT NON_TER 120
SQ SEQUENCE 120 AA; 13600 MW; EECADA4FB9B9639B CRC64;

Query Match 100.0%; Score 25; DB 8; Length 120;
Best Local Similarity 100.0%; Pred. No. 96;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
Db 60 SSVMH 64
|||||

RESULT 11
Q9TB72 ID Q9TB72 PRELIMINARY; PRT; 128 AA.
AC Q9TB72;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Cytochrome b (Fragment).
GN CYTB.
OS Lampetra fluviatilis (River lamprey).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Lampetra.
OX NCBI_TaxID=7748;
RN [1]
RP SEQUENCE FROM N.A.
RA Döcker M.F., Youson J.H., Beamish R.J., Devlin R.H.;
RT "Phylogeny of the lamprey genus Lampetra inferred from mitochondrial
RT cytochrome b and ND3 gene sequences.";
RL Can. J. Fish. Aquat. Sci. 56:2340-2349(1999).
CC -!- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
CC COUPLED TO ATP SYNTHESIS (BY SIMILARITY).
CC -!- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
CC BOUND TO THE PROTEIN (BY SIMILARITY).
CC -!- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
CC CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.

```

CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
 DR EMBL; AF177949; AAF00996.1; --
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO; GO:0005739; C:mitochondrion; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR GO; GO:0005810; P:transport; IEA.
 DR InterPro; IPR005797; Cytochrome b N.
 DR Pfam; PF00033; cytochrome b N; 1.
 DR PROSITE; PS00192; CYTOCHROME B HEME; 1.
 KW Electron transport; Heme; Respiratory chain; Transmembrane; Transport;
 KW Mitochondrion.
 FT NON_TER 128 128
 SQ SEQUENCE 128 AA; 14301 MW; B8CCB1B81EEC494A CRC64;
 Query Match 100.0%; Score 25; DB 8; Length 128;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SSVMH 5
 DB 68 SSVMH 72
 RESULT 12
 ID Q8MIH7 PRELIMINARY; PRT; 128 AA.
 AC Q8MIH7;
 DT 01-OCT-2002 (Tremblrel. 22, Created)
 DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
 DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
 DE Cytochrome b (Fragment).
 OS Mesopodion peruvianus (Peruvian beaked whale).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Cetacea; Odontoceti; Ziphiidae;
 OC Mesopodion.
 OX NCBI_TaxID=27617;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Van Heiden A.L., Baker A.N., Dalebout M.L., Reyes J.C.,
 RA Van Waerebeek K., Baker C.S.;
 RT "Resurrection of Mesopodion traversii (Gray, 1874), senior synonym of
 RT M. bahamondi Reyes, Van Waerebeek, Cardenas and Yanez, 1995 (Cetacea;
 RT Ziphiidae).";
 RL Mar. Mamm. Sci. 18:0-0(2002).
 CC -1- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
 CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
 CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
 CC COUPLED TO ATP SYNTHESIS (BY SIMILARITY).
 CC -1- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
 CC BOUND TO THE PROTEIN (BY SIMILARITY).
 CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
 CC CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
 DR EMBL; AF492414; AAW74073.1; --
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO; GO:0005739; C:mitochondrion; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR GO; GO:0005810; P:transport; IEA.
 DR InterPro; IPR005797; Cytochrome b N.
 DR Pfam; PF00033; cytochrome b N; 1.
 DR PROSITE; PS00192; CYTOCHROME B HEME; 1.
 KW Electron transport; Heme; Respiratory chain; Transmembrane; Transport;
 KW Mitochondrion.
 FT NON_TER 128 128
 SQ SEQUENCE 128 AA; 14611 MW; 5BF9B08C2A6CC991 CRC64;
 Query Match 100.0%; Score 25; DB 8; Length 128;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SSVMH 5
 DB 64 SSVMH 68
 RESULT 13
 ID Q9T3E3 PRELIMINARY; PRT; 128 AA.
 AC Q9T3E3;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
 DE Cytochrome B (Fragment).
 GN CYTB.
 OS Entosphenus similis.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
 OC Petromyzontiformes; Petromyzontidae; Entosphenus.
 OX NCBI_TaxID=104636;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ORE 07, CAL 01, and ORE 01;
 RA Docker M.F., Youson J.H., Beamish R.J., Devlin R.H.;
 RT "Phylogeny of the lamprey genus Lampetra inferred from mitochondrial
 RT cytochrome b and ND3 gene sequences";
 RL Can. J. Fish. Aquat. Sci. 0:0-0(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Lorton C.M., Markle D.F., Reid S.B., Docker M.F.;
 RT "Redescription of the presumed-extinct Miller Lake lamprey, Lampetra
 RT minima";
 RL Copeia 2000:1019-1028(2000).
 CC -1- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
 CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
 CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
 CC COUPLED TO ATP SYNTHESIS (BY SIMILARITY).
 CC -1- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
 CC BOUND TO THE PROTEIN (BY SIMILARITY).
 CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
 CC CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
 DR EMBL; AF177957; AAF01004.1; --
 DR EMBL; AF177955; AAF01002.1; --
 DR EMBL; AF177956; AAF01003.1; --
 DR EMBL; AF257134; AAG13284.1; --
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO; GO:0005739; C:mitochondrion; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR GO; GO:0005810; P:transport; IEA.
 DR InterPro; IPR005797; Cytochrome b N.
 DR Pfam; PF00033; cytochrome b N; 1.
 DR PROSITE; PS00192; CYTOCHROME B HEME; 1.
 KW Electron transport; Heme; Respiratory chain; Transmembrane; Transport;
 KW Mitochondrion.
 FT NON_TER 128 128
 SQ SEQUENCE 128 AA; 14301 MW; BABCB4B81EEC4C4A CRC64;
 Query Match 100.0%; Score 25; DB 8; Length 128;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
 DB 68 SSVMH 72
 RESULT 14
 Q9TB73

ID Q9TB73 PRELIMINARY; PRT; 128 AA.
 AC Q9TB73;
 DT 01-MAY-2000 (TREMELrel. 13, Created)
 DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
 DE Cytochrome b (Fragment).
 GN CYTB.
 OS Lampetra richardsoni (Western brook lamprey).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
 OC Petromyzontiformes; Petromyzontidae; Lampetra.
 OX NCBI_TaxID=104635;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Docker M.F., Youson J.H., Beamish R.J., Devlin R.H.;
 RT "Phylogeny of the lamprey genus Lampetra inferred from mitochondrial
 cytochrome b and ND3 gene sequences.";
 RL Can. J. Fish. Aquat. Sci. 56:2340-2349(1999).
 CC -!- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
 COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
 RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
 COUPLED TO ATP SYNTHESIS (BY SIMILARITY).
 CC -!- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
 BOUND TO THE PROTEIN (BY SIMILARITY).
 CC -!- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
 CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
 DR EMBL; AF177948; AAF00995.1; -;
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO; GO:0005739; C:mitochondrion; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR005797; Cytb_b6_N.
 DR Pfam; PF00033; cytochrome_b_N; 1.
 DR PROSITE; PS00192; CYTOCHROME_B_HEME; 1.
 KW Electron transport; Heme; Respiratory chain; Transmembrane; Transport;
 KW Mitochondrion.
 FT NON_TER 128
 SQ SEQUENCE 128 AA; 14301 MW; BABCB4B81EEC4C4A CRC64;
 Query Match 100.0%; Score 25; DB 8; Length 128;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SSVMH 5
 DB 68 SSVMH 72
 RESULT 15
 Q9TB74 PRELIMINARY; PRT; 128 AA.
 ID Q9TB74;
 DT 01-MAY-2000 (TREMELrel. 13, Created)
 DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
 DE Cytochrome b (Fragment).
 GN CYTB.
 OS Lampetra ayresi (River lamprey).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
 OC Petromyzontiformes; Petromyzontidae; Lampetra.
 OX NCBI_TaxID=104631;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Docker M.F., Youson J.H., Beamish R.J., Devlin R.H.;
 RT "Phylogeny of the lamprey genus Lampetra inferred from mitochondrial
 cytochrome b and ND3 gene sequences.";
 RL Can. J. Fish. Aquat. Sci. 56:2340-2349(1999).
 CC -!- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
 COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A

CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
 COUPLED TO ATP SYNTHESIS (BY SIMILARITY).
 CC COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
 BOUND TO THE PROTEIN (BY SIMILARITY).
 CC -!- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
 CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
 DR EMBL; AF177947; AAF00994.1; -;
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO; GO:0005739; C:mitochondrion; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR005797; Cytb_b6_N.
 DR Pfam; PF00033; cytochrome_b_N; 1.
 DR PROSITE; PS00192; CYTOCHROME_B_HEME; 1.
 KW Electron transport; Heme; Respiratory chain; Transmembrane; Transport;
 KW Mitochondrion.
 FT NON_TER 128
 SQ SEQUENCE 128 AA; 14301 MW; BABCB4B81EEC4C4A CRC64;
 Query Match 100.0%; Score 25; DB 8; Length 128;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SSVMH 5
 DB 68 SSVMH 72
 RESULT 16
 Q9TB69 PRELIMINARY; PRT; 128 AA.
 ID Q9TB69;
 DT 01-MAY-2000 (TREMELrel. 13, Created)
 DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)
 DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
 DE Cytochrome B (Fragment).
 GN CYTB.
 OS Entosphenus lethophagus.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
 OC Petromyzontiformes; Petromyzontidae; Entosphenus.
 OX NCBI_TaxID=104633;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Docker M.F., Youson J.H., Beamish R.J., Devlin R.H.;
 RT "Phylogeny of the lamprey genus Lampetra inferred from mitochondrial
 cytochrome b and ND3 gene sequences.";
 RL Can. J. Fish. Aquat. Sci. 0:0-0(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Lorion C.M., Markle D.F., Reid S.B., Docker M.F.;
 RT "Redescription of the presumed-extinct Miller Lake lamprey, Lampetra
 minima.";
 RL Copeia 2000:1019-1028(2000).
 CC -!- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
 COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
 RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
 COUPLED TO ATP SYNTHESIS (BY SIMILARITY).
 CC COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
 BOUND TO THE PROTEIN (BY SIMILARITY).
 CC -!- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
 CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
 DR EMBL; AF177954; AAF01001.1; -;
 DR EMBL; AF257129; AAG13279.1; -;
 DR EMBL; AF257124; AAG13275.1; -;
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO; GO:0005739; C:mitochondrion; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.

DR GO:0006118; P:electron transport; IEA.
 DR GO:0006810; P:transport; IEA.
 DR InterPro: IPR005797; Cytochrome b N.
 DR Pfam: PF00033; cytochrome b N; 1.
 DR PROSITE: PS00192; CYTOCHROME B HEME; 1.
 KW Electron transport; Heme; Respiratory chain; Transmembrane; Transport;
 KW Mitochondrion.
 FT NON_TER 128
 SQ SEQUENCE 128 AA; 14301 MW; BABCB4B81BEC4C4A CRC64;
 Query Match 100.0%; Score 25; DB 8; Length 128;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SSVMH 5
 DB 68 SSVMH 72
 RESULT 17
 Q9G155 PRELIMINARY; PRT; 128 AA.
 AC Q9G155;
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
 DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
 DE Cytochrome b (fragment).
 GN CYTB.
 OS Entosphenus minimus.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
 OC Petromyzontiformes; Petromyzontidae; Entosphenus.
 OX NCBI_TaxID=126726;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lorion C.M., Markle D.F., Reid S.B., Docker M.F.;
 RT "Redescription of the presumed-extinct Miller Lake lamprey, Lampetra minima";
 RL Copeia 2000:1019-1028(2000).
 CC -!- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL COUPLED TO ATP SYNTHESIS (BY SIMILARITY).
 CC -!- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY BOUND TO THE PROTEIN (BY SIMILARITY).
 CC -!- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B, CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
 DR EMBL: AF257132; AAG13282.1; -;
 DR EMBL: AF257130; AAG13280.1; -;
 DR EMBL: AF257131; AAG13281.1; -;
 DR GO:00016021; C:integral to membrane; IEA.
 DR GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO:0005739; C:mitochondrion; IEA.
 DR GO:0016491; F:oxidoreductase activity; IEA.
 DR GO:0006118; P:electron transport; IEA.
 DR GO:0006810; P:transport; IEA.
 DR InterPro: IPR005797; Cytochrome b N.
 DR Pfam: PF00033; cytochrome b N; 1.
 DR PROSITE: PS00192; CYTOCHROME B HEME; 1.
 KW Electron transport; Heme; Respiratory chain; Transmembrane; Transport;
 KW Mitochondrion.
 FT NON_TER 128
 SQ SEQUENCE 128 AA; 14301 MW; BABCB4B81BEC4C4A CRC64;
 Query Match 100.0%; Score 25; DB 8; Length 128;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SSVMH 5
 DB 68 SSVMH 72

RESULT 18
 Q9T656 PRELIMINARY; PRT; 128 AA.
 AC Q9T656;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
 DE Cytochrome b (fragment).
 GN CYTB.
 OS Lampetra japonica (Japanese lamprey) (Entosphenus japonicus).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
 OC Petromyzontiformes; Petromyzontidae; Lethenteron.
 OX NCBI_TaxID=94989;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Docker M.F., Youson J.H., Beamish R.J., Devlin R.H.;
 RT "Phylogeny of the lamprey genus Lampetra inferred from mitochondrial cytochrome b and ND3 gene sequences";
 RL Can. J. Fish. Aquat. Sci. 56:2340-2349(1999).
 CC -!- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL COUPLED TO ATP SYNTHESIS (BY SIMILARITY).
 CC -!- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY BOUND TO THE PROTEIN (BY SIMILARITY).
 CC -!- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B, CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
 DR EMBL: AF177959; AAF01006.1; -;
 DR GO:00016021; C:integral to membrane; IEA.
 DR GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO:0005739; C:mitochondrion; IEA.
 DR GO:0016491; F:oxidoreductase activity; IEA.
 DR GO:0006118; P:electron transport; IEA.
 DR GO:0006810; P:transport; IEA.
 DR InterPro: IPR005797; Cytochrome b N.
 DR Pfam: PF00033; cytochrome b N; 1.
 DR PROSITE: PS00192; CYTOCHROME B HEME; 1.
 KW Electron transport; Heme; Respiratory chain; Transmembrane; Transport;
 KW Mitochondrion.
 FT NON_TER 128
 SQ SEQUENCE 128 AA; 14314 MW; AF5CB5804EC4C4A CRC64;
 Query Match 100.0%; Score 25; DB 8; Length 128;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SSVMH 5
 DB 68 SSVMH 72
 RESULT 19
 Q9G9U6 PRELIMINARY; PRT; 128 AA.
 AC Q9G9U6;
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
 DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
 DE Cytochrome b (fragment).
 GN CYTB.
 OS Entosphenus cf. similis.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
 OC Petromyzontiformes; Petromyzontidae; Entosphenus.
 OX NCBI_TaxID=128046;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lorion C.M., Markle D.F., Reid S.B., Docker M.F.;
 RT "Redescription of the presumed-extinct Miller Lake lamprey, Lampetra minima";

RL Copeia 2000:1019-1028 (2000).
 CC -1- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
 CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
 CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
 CC COUPLED TO ATP SYNTHESIS (BY SIMILARITY).
 CC -1- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
 CC BOUND TO THE PROTEIN (BY SIMILARITY).
 CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
 CC CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
 DR EMBL; AF257128; AG13278.1; --
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO; GO:0005739; C:mitochondrion; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR GO; GO:0006810; P:transpot; IEA.
 DR InterPro; IPR005797; Cytb_b6_N.
 DR Pfam; PF00033; cytochrome b_N; 1.
 DR PROSITE; PS00192; CYTOCHROME B_HEME; 1.
 KW Electron transport; Heme; Respiratory chain; Transmembrane; Transport;
 KW Mitochondrion.
 FT NON_TER 128 128
 SQ SEQUENCE 128 AA; 14301 MW; BABCB4B81EEC4C4A CRC64;
 Query Match 100.0%; Score 25; DB 8; Length 128;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SSVMH 5
 Db 68 SSVMH 72
 RESULT 20
 Q9TB66 PRELIMINARY; PRT; 128 AA.
 AC Q9TB66;
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
 DE Cytochrome b (Fragment).
 GN CYTB.
 OS Petromyzon marinus (Sea lamprey).
 OC Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
 OC Petromyzontiformes; Petromyzontidae; Petromyzon.
 OX NCBI_TaxID=7757;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Docker M.F., Youson J.H., Beamish R.J., Devlin R.H.;
 RT "Phylogeny of the lamprey genus Lampetra inferred from mitochondrial
 RT cytochrome b and ND3 gene sequences";
 RL Can. J. Fish. Aquat. Sci. 56:2340-2349 (1999).
 CC -1- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
 CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
 CC COUPLED TO ATP SYNTHESIS (BY SIMILARITY).
 CC -1- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
 CC BOUND TO THE PROTEIN (BY SIMILARITY).
 CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
 CC CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
 DR EMBL; AF177953; AAF01008.1; --
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO; GO:0005739; C:mitochondrion; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR GO; GO:0006810; P:transpot; IEA.
 DR InterPro; IPR005797; Cytb_b6_N.
 DR Pfam; PF00033; cytochrome b_N; 1.
 DR PROSITE; PS00192; CYTOCHROME B_HEME; 1.
 KW Electron transport; Heme; Respiratory chain; Transmembrane; Transport;
 KW Mitochondrion.
 FT NON_TER 128 128
 SQ SEQUENCE 128 AA; 14301 MW; BABCB4B81EEC4C4A CRC64;
 Query Match 100.0%; Score 25; DB 8; Length 128;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SSVMH 5
 Db 68 SSVMH 72

KW Electron transport; Heme; Respiratory chain; Transmembrane; Transport;
 KW Mitochondrion.
 FT NON_TER 128 128
 SQ SEQUENCE 128 AA; 14318 MW; BEACB618195C4D4A CRC64;
 Query Match 100.0%; Score 25; DB 8; Length 128;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SSVMH 5
 Db 68 SSVMH 72
 RESULT 21
 Q9TB70 PRELIMINARY; PRT; 128 AA.
 AC Q9TB70;
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
 DE Cytochrome b (Fragment).
 GN CYTB.
 OS Entosphenus macrostomus.
 OC Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
 OC Petromyzontiformes; Petromyzontidae; Entosphenus.
 OX NCBI_TaxID=104634;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Docker M.F., Youson J.H., Beamish R.J., Devlin R.H.;
 RT "Phylogeny of the lamprey genus Lampetra inferred from mitochondrial
 RT cytochrome b and ND3 gene sequences";
 RL Can. J. Fish. Aquat. Sci. 56:2340-2349 (1999).
 CC -1- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
 CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
 CC COUPLED TO ATP SYNTHESIS (BY SIMILARITY).
 CC -1- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
 CC BOUND TO THE PROTEIN (BY SIMILARITY).
 CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
 CC CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
 DR EMBL; AF177953; AAF01000.1; --
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO; GO:0005739; C:mitochondrion; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR GO; GO:0006810; P:transpot; IEA.
 DR InterPro; IPR005797; Cytb_b6_N.
 DR Pfam; PF00033; cytochrome b_N; 1.
 DR PROSITE; PS00192; CYTOCHROME B_HEME; 1.
 KW Electron transport; Heme; Respiratory chain; Transmembrane; Transport;
 KW Mitochondrion.
 FT NON_TER 128 128
 SQ SEQUENCE 128 AA; 14301 MW; BABCB4B81EEC4C4A CRC64;
 Query Match 100.0%; Score 25; DB 8; Length 128;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SSVMH 5
 Db 68 SSVMH 72
 RESULT 22
 Q9TB71 PRELIMINARY; PRT; 128 AA.
 AC Q9TB71;
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)

DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Cytochrome b (Fragment).
GN Cytb.
OS Lampetra tridentata (Pacific lamprey) (Entosphenus tridentatus).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Entosphenus.
OX NCBI_TaxID=7751;
RN [1]
RN SEQUENCE FROM N.A.
RA Docker M.F., Youson J.H., Beamish R.J., Devlin R.H.;
RP "Phylogeny of the lamprey genus Lampetra inferred from mitochondrial
RT cytochrome b and ND3 gene sequences";
RL Can. J. Fish. Aquat. Sci. 56:2340-2349 (1999).
CC -!- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
CC COUPLED TO ATP SYNTHESIS (BY SIMILARITY).
CC -!- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
CC BOUND TO THE PROTEIN (BY SIMILARITY).
CC -!- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
CC CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
DR EMBL; AF177958; AAF01005.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005797; Cytb_b6_N.
DR Pfam; PF00033; cytochrome_b_N; 1.
KW Electron transport; Heme; Respiratory chain; Transmembrane; Transport;
KW Mitochondrion.
FT NON_TER 128 128
SQ SEQUENCE 128 AA; 14301 MW; BABCB4B81EEC4C4A CRC64;
Query Match 100.0%; Score 25; DB 8; Length 128;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SSVMH 5
DB 68 SSVMH 72
RESULT 23
Q9TB67 PRELIMINARY; PRT; 128 AA.
AC Q9TB67;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Cytochrome b (Fragment).
GN Cytb.
OS Entosphenus hubbsi (Kern brook lamprey).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Entosphenus.
OX NCBI_TaxID=104630;
RN [1]
RN SEQUENCE FROM N.A.
RA Docker M.F., Youson J.H., Beamish R.J., Devlin R.H.;
RP "Phylogeny of the lamprey genus Lampetra inferred from mitochondrial
RT cytochrome b and ND3 gene sequences";
RL Can. J. Fish. Aquat. Sci. 56:2340-2349 (1999).
CC -!- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
CC COUPLED TO ATP SYNTHESIS (BY SIMILARITY).
CC -!- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
CC BOUND TO THE PROTEIN (BY SIMILARITY).
CC -!- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
CC CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
DR EMBL; AF177958; AAF01005.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0005739; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR005797; Cytb_b6_N.
DR Pfam; PF00033; cytochrome_b_N; 1.
DR ProSite; PS00192; CYTOCHROME_B_HEME; 1.
KW Electron transport; Heme; Respiratory chain; Transmembrane; Transport;
KW Mitochondrion.
FT NON_TER 128 128
SQ SEQUENCE 128 AA; 14301 MW; BABCB4B81EEC4C4A CRC64;
Query Match 100.0%; Score 25; DB 8; Length 128;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SSVMH 5
DB 68 SSVMH 72
RESULT 24
Q9TB68 PRELIMINARY; PRT; 128 AA.
AC Q9TB68;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Cytochrome b (Fragment).
GN Cytb.
OS Entosphenus hubbsi (Kern brook lamprey).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Entosphenus.
OX NCBI_TaxID=104632;
RN [1]
RN SEQUENCE FROM N.A.
RA Docker M.F., Youson J.H., Beamish R.J., Devlin R.H.;
RP "Phylogeny of the lamprey genus Lampetra inferred from mitochondrial
RT cytochrome b and ND3 gene sequences";
RL Can. J. Fish. Aquat. Sci. 56:2340-2349 (1999).
CC -!- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
CC COUPLED TO ATP SYNTHESIS (BY SIMILARITY).
CC -!- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
CC BOUND TO THE PROTEIN (BY SIMILARITY).
CC -!- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
CC CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
DR EMBL; AF177958; AAF01005.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0005739; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.

[1]
RN SEQUENCE FROM N.A.
RP Docker M.F., Youson J.H., Beamish R.J., Devlin R.H.;
RA "Phylogeny of the lamprey genus Lampetra inferred from mitochondrial
RT cytochrome b and ND3 gene sequences";
RL Can. J. Fish. Aquat. Sci. 56:2340-2349 (1999).
CC -!- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
CC COUPLED TO ATP SYNTHESIS (BY SIMILARITY).
CC -!- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
CC BOUND TO THE PROTEIN (BY SIMILARITY).
CC -!- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
CC CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
DR EMBL; AF177960; AAF01007.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005797; Cytb_b6_N.
DR Pfam; PF00033; cytochrome_b_N; 1.
KW Electron transport; Heme; Respiratory chain; Transmembrane; Transport;
KW Mitochondrion.
FT NON_TER 128 128
SQ SEQUENCE 128 AA; 14386 MW; ACD10D80B4EC4C4A CRC64;
Query Match 100.0%; Score 25; DB 8; Length 128;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SSVMH 5
DB 68 SSVMH 72
RESULT 24
Q9TB68 PRELIMINARY; PRT; 128 AA.
AC Q9TB68;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Cytochrome b (Fragment).
GN Cytb.
OS Entosphenus hubbsi (Kern brook lamprey).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Entosphenus.
OX NCBI_TaxID=104632;
RN [1]
RN SEQUENCE FROM N.A.
RA Docker M.F., Youson J.H., Beamish R.J., Devlin R.H.;
RP "Phylogeny of the lamprey genus Lampetra inferred from mitochondrial
RT cytochrome b and ND3 gene sequences";
RL Can. J. Fish. Aquat. Sci. 56:2340-2349 (1999).
CC -!- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
CC COUPLED TO ATP SYNTHESIS (BY SIMILARITY).
CC -!- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
CC BOUND TO THE PROTEIN (BY SIMILARITY).
CC -!- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
CC CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
DR EMBL; AF177958; AAF01005.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0005739; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.

DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005797; Cytb b6 N.
DR Pfam; PF00033; cytochrome_b6; 1.
DR PROSITE; PS00192; CYTOCHROME_B_HEME; 1.
KW Electron transport; Heme; Respiratory chain; Transmembrane; Transport;
KW Mitochondrion.
FT NON_TER 128
SQ SEQUENCE 128 AA; 14287 MW; 40BCA119C947802D CRC64;

Query Match 100.0%; Score 25; DB 8; Length 128;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
DB 68 SSVMH 72

RESULT 25
O21661
ID O21661 PRELIMINARY; PRT; 185 AA.
AC O21661;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cytochrome b (Fragment).
GN CYTB.
OS Hydromantes brunus.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Caudata; Salamandroidea; Plethodontidae;
OC Hydromantes.
OX NCBI_TaxID=57553;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97398713; PubMed=9254927;
RA Jackman T.R., Applebaum G., Wake D.B.;
RT "Phylogenetic relationships of bolitoglossine salamanders: a
RT demonstration of the effects of combining morphological and molecular
RT data sets".
RL Mol. Biol. Evol. 14:883-891(1997).
CC -!- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
CC COUPLED TO ATP SYNTHESIS (BY SIMILARITY).
CC -!- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
CC BOUND TO THE PROTEIN (BY SIMILARITY).
CC -!- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
CC CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
DR EMBL; U99614; AAB68635.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005797; Cytb b6 N.
DR Pfam; PF00033; cytochrome_b6; 1.
DR PROSITE; PS00192; CYTOCHROME_B_HEME; 1.
KW Electron transport; Heme; Respiratory chain; Transmembrane; Transport;
KW Mitochondrion.
FT NON_TER 1 185
FT NON_TER 185 185
SQ SEQUENCE 185 AA; 20851 MW; 471DB2B09ECF251A CRC64;

Query Match 100.0%; Score 25; DB 8; Length 185;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
DB 51 SSVMH 55

Search completed: September 24, 2004, 02:18:40
Job time : 30.9623 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2005, 17:57:45 ; Search time 32.3438 Seconds
(without alignments)
79.162 Million cell updates/sec

Title: US-10-088-639A-2_COPY_158_162
Perfect score: 25
Sequence: 1 SSVNH 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*

1: uniprot_sprot:*

2: uniprot_trembl:*

SUMMARIES

Result No.	Query Match	Length	DB	ID	Description
1	25	100.0	78	Q8VNL9	Q8VNL9 pseudomonas
2	25	100.0	90	Q9PW02	Q9PW02 petromyzon
3	25	100.0	90	Q9Y072	Q9Y072 lampetra pl
4	25	100.0	93	Q35383	Q35383 pantodon bu
5	25	100.0	100	Q9Y347	Q9Y347 homo sapien
6	25	100.0	102	Q9B378	Q9B378 dicologloss
7	25	100.0	102	Q9XLK3	Q9XLK3 buglossidii
8	25	100.0	110	Q9T9U6	Q9T9U6 puntius con
9	25	100.0	120	Q99070	Q99070 fundulosoma
10	25	100.0	120	Q99071	Q99071 pronothobra
11	25	100.0	122	Q64GF1	Q64GF1 plethodon e
12	25	100.0	128	Q8M1H7	Q8M1H7 mesoplodon
13	25	100.0	128	Q6Y4C8	Q6Y4C8 plethodon e
14	25	100.0	128	Q6Y4D1	Q6Y4D1 plethodon e
15	25	100.0	128	Q9G155	Q9G155 entosphenus
16	25	100.0	128	Q9G9U6	Q9G9U6 entosphenus
17	25	100.0	128	Q9T3E3	Q9T3E3 entosphenus
18	25	100.0	128	Q9TB66	Q9TB66 lampetra ja
19	25	100.0	128	Q9TB66	Q9TB66 petromyzon
20	25	100.0	128	Q9TB67	Q9TB67 lampetra ap
21	25	100.0	128	Q9TB68	Q9TB68 entosphenus
22	25	100.0	128	Q9TB69	Q9TB69 entosphenus
23	25	100.0	128	Q9TB70	Q9TB70 entosphenus
24	25	100.0	128	Q9TB71	Q9TB71 lampetra tr
25	25	100.0	128	Q9TB72	Q9TB72 lampetra fl
26	25	100.0	128	Q9TB73	Q9TB73 lampetra ri
27	25	100.0	128	Q9TB74	Q9TB74 lampetra ay
28	25	100.0	133	Q9EGG7	Q9EGG7 orthogomys
29	25	100.0	140	Q9G7S1	Q9G7S1 coetomys me
30	25	100.0	185	Q21661	Q21661 hydronantes
31	25	100.0	185	Q7GF68	Q7GF68 aneides lug

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Query Match 100.0% Score 25; DB 2; Length 78;
Best Local Similarity 100.0%; Pred.No. 61;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSVNH 5
Db 59 SSVNH 63

RESULT 2

Q9PW02
ID Q9PW02 PRELIMINARY; PRT; 90 AA.
AC Q9PW02;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Cytochrome b (Fragment).
GN Name=cyt b;
OS Petromyzon marinus (Sea lamprey).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Hypercoartia;
OC Petromyzontiformes; Petromyzontidae; Petromyzon.
OX NCBI_TaxID=7757;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RA Tagliavini J., Tizi R., Conterio F., Mariottini P., Gandolfi G.;
RT "Mitochondrial DNA sequences in three genera of Italian lampreys."
RL Boll. Zool. 61:331-333(1994).
DR EMBL; X79111; CAA55728.2; -;
GO; GO:0016020; C:membrane; IEA.

ALIGNMENTS

RESULT 1

Q8VNL9
ID Q8VNL9 PRELIMINARY; PRT; 78 AA.
AC Q8VNL9;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Pseudomonas putida.
OG Plasmid pMW0.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22423065; PubMed=12534468;
RA Greated A., Lamberton L., Williams P.A., Thomas C.M.;
RT "Complete sequence of the IncP-9 TOL plasmid pMW0 from Pseudomonas putida."
RL Environ. Microbiol. 4:856-871(2002).
DR EMBL; AJ344068; CAC86770.1; -;
SQ Hypothetical protein; Plasmid.
KW SEQUENCE 78 AA; 8805 MW; 8762D76ACESFF30A CRC64;

32 25 100.0 189 2 Q7V858
33 25 100.0 214 2 Q8RJZ9
34 25 100.0 217 2 Q6W236
35 25 100.0 221 2 Q8UCT3
36 25 100.0 222 2 Q33739
37 25 100.0 224 2 Q6SMS2
38 25 100.0 224 2 Q6SMW8
39 25 100.0 224 2 Q6SMZ4
40 25 100.0 224 2 Q6SN31
41 25 100.0 224 2 Q6SN86
42 25 100.0 224 2 Q6SND0
43 25 100.0 224 2 Q6SNG2
44 25 100.0 224 2 Q6SNL5
45 25 100.0 244 1 CAH_ERWCA

Q7V858 prochloroco
Q8RJZ9 rhizobium l
Q6W236 rhizobium s
Q8UCT3 agrobacteri
Q33739 aneides lug
Q6SMS2 lagothrix l
Q6SMW8 ates azara
Q6SMZ4 actus azara
Q6SN31 cercopithe
Q6SN86 eulemur mon
Q6SND0 alouatta ca
Q6SNG2 hylobates s
Q6SNL5 cebus apell
Q52538 erwinia car

|||||
43 SSVMH 47

RESULT 6

Q9B378 PRELIMINARY; PRT; 102 AA.
AC Q9B378;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cytochrome b (Fragment).
GN Name=cytb;
OS Dicotloglossa cuneata.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
OC Soleiidae; Soleidae; Dicotloglossa.
OX NCBI_TaxID=153209;
RN [1]
RP SEQUENCE FROM N.A.
RA Cunha R., Caetano S., Castilho R.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Component of the ubiquinol-cytochrome c reductase complex (complex III or cytochrome b-c1 complex), which is a respiratory chain that generates an electrochemical potential coupled to ATP synthesis (By similarity).
CC -!- COFACTOR: Binds two heme groups non-covalently. Heme 1 (or BL or b562) is low-potential and absorbs at about 562 nm, and heme 2 (or BH or b566) is high-potential and absorbs at about 566 nm (By similarity).
CC -!- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b, cytochrome c1 and the Rieske protein (By similarity).
CC -!- SIMILARITY: Belongs to the cytochrome b family.
DR EMBL; AF338429; AAK30100.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005797; Cytb_b6_N.
DR Pfam; PF00033; Cytochrom_B_N; 1.
DR PROSITE; PS00192; CYTOCHROME_B_HEME; 1.
DR PROSITE; PS00442; GATASE_TYPE_I; UNKNOWN 1.
KW Electron transport; Heme; Mitochondrion; Respiratory chain; Transmembrane; Transport.
FT NON_TER 1
FT NON_TER 102 102
SQ SEQUENCE 102 AA; 11318 MW; C89116AD5009B699 CRC64;

Query Match 100.0%; Score 25; DB 2; Length 102;
Best Local Similarity 100.0%; Pred. NO. 82;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
|||||
DB 32 SSVMH 36

RESULT 7

Q9XLK3 PRELIMINARY; PRT; 102 AA.
AC Q9XLK3;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cytochrome b (Fragment).
GN Name=cytb;
OS Buglossidium luteum.
OG Mitochondrion.

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
OC Soleiidae; Soleidae; Buglossidium.
OX NCBI_TaxID=90071;
RN [1]
RP SEQUENCE FROM N.A.
RA Tinti F., Piccinetti C., Tommasini S., Vallisneri M.;
RT Mitochondrial DNA Variation, Phylogenetic Relationships, and Evolution of Four Mediterranean Genera of Soles (Soleidae, Pleuronectiformes).
RL Mar. Biotechnol. 2:274-284 (2000).
CC -!- FUNCTION: Component of the ubiquinol-cytochrome c reductase complex (complex III or cytochrome b-c1 complex), which is a respiratory chain that generates an electrochemical potential coupled to ATP synthesis (By similarity).
CC -!- COFACTOR: Binds two heme groups non-covalently. Heme 1 (or BL or b562) is low-potential and absorbs at about 562 nm, and heme 2 (or BH or b566) is high-potential and absorbs at about 566 nm (By similarity).
CC -!- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b, cytochrome c1 and the Rieske protein (By similarity).
CC -!- SIMILARITY: Belongs to the cytochrome b family.
DR EMBL; AF113206; AAD23435.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005797; Cytb_b6_N.
DR Pfam; PF00033; Cytochrom_B_N; 1.
DR PROSITE; PS00192; CYTOCHROME_B_HEME; 1.
KW Electron transport; Heme; Mitochondrion; Respiratory chain; Transmembrane; Transport.
FT NON_TER 1
FT NON_TER 102 102
SQ SEQUENCE 102 AA; 11356 MW; 90F4444E2DD2B4E3 CRC64;
Query Match 100.0%; Score 25; DB 2; Length 102;
Best Local Similarity 100.0%; Pred. NO. 82;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
|||||
DB 32 SSVMH 36

RESULT 8
Q9T9U6 PRELIMINARY; PRT; 110 AA.
AC Q9T9U6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cytochrome b (Fragment).
GN Name=cyt b;
OS Puntius conchionius (Rosy barb).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Puntius.
OX NCBI_TaxID=27708;
RN [1]
RP SEQUENCE FROM N.A.
RA Gilles A., Lecoindre G.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Component of the ubiquinol-cytochrome c reductase complex (complex III or cytochrome b-c1 complex), which is a respiratory chain that generates an electrochemical potential coupled to ATP synthesis (By similarity).
CC -!- COFACTOR: Binds two heme groups non-covalently. Heme 1 (or BL or

b562) is low-potential and absorbs at about 562 nm, and heme 2 (or BH or b566) is high-potential and absorbs at about 566 nm (By similarity).
 -!- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b, cytochrome c1 and the Rieske protein (By similarity).
 -!- SIMILARITY: Belongs to the cytochrome b family.
 DR EMBL; AJ388455; CAB62077.1; -.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO; GO:0005739; C:mitochondrion; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR005797; Cytb_b6_N.
 DR Pfam; PF00033; Cytochrom_B_N; 1.
 DR PROSITE; PS00192; CYTOCHROME_B_HEME; 1.
 KW Electron transport; Heme; Mitochondrion; Respiratory chain; Transmembrane; Transport.
 FT NON_TER 1
 FT NON_TER 110 110
 SQ SEQUENCE 110 AA; 12598 MW; FAA7FE8EDDA70AC3 CRC64;
 Query Match 100.0%; Score 25; DB 2; Length 110;
 Best Local Similarity 100.0%; Pred. No. 89;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SSVMH 5
 Db 33 SSVMH 37
 RESULT 9
 Q99070 PRELIMINARY; PRT; 120 AA.
 AC O99070;
 DT 01-MAY-1999 (TReMBLrel. 10, Created)
 DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
 DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
 DE Cytochrome b (Fragment).
 GN Name=cytb;
 OS Fundulosoma thieryi.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha; OC Cyprinodontiformes; Aplocheilidae; Aplocheilinae; Fundulosoma. OX NCBI_TaxID=60304;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Aquarium strain;
 RX MEDLINE=99214037; PubMed=10196077; DOI=10.1006/mpev.1998.0566;
 RA Murphy W.J., Collier G.E.;
 RT "Phylogenetic relationships of African killifishes in the genera Aphyosemon and Fundulopanchax inferred from mitochondrial DNA sequences.";
 RL Mol. Phylogenet. Evol. 11:351-360(1999).
 CC -!- FUNCTION: Component of the ubiquinol-cytochrome c reductase complex (complex III or cytochrome b-c1 complex), which is a respiratory chain that generates an electrochemical potential coupled to ATP synthesis (By similarity).
 CC -!- COPACTOR: Binds two heme groups non-covalently. Heme 1 (or BL or b562) is low-potential and absorbs at about 562 nm, and heme 2 (or BH or b566) is high-potential and absorbs at about 566 nm (By similarity).
 CC -!- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b, cytochrome c1 and the Rieske protein (By similarity).
 CC -!- SIMILARITY: Belongs to the cytochrome b family.
 DR EMBL; AF002284; RAD00945.1; -.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO; GO:0005739; C:mitochondrion; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR005797; Cytb_b6_N.
 DR Pfam; PF00033; Cytochrom_B_N; 1.
 DR PROSITE; PS00192; CYTOCHROME_B_HEME; 1.
 KW Electron transport; Heme; Mitochondrion; Respiratory chain; Transmembrane; Transport.
 FT NON_TER 1
 FT NON_TER 120 120
 SQ SEQUENCE 120 AA; 13591 MW; 54D1FD7D50C1538CE CRC64;
 Query Match 100.0%; Score 25; DB 2; Length 120;
 Best Local Similarity 100.0%; Pred. No. 98;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SSVMH 5
 Db 60 SSVMH 64
 RESULT 10
 Q99071 PRELIMINARY; PRT; 120 AA.
 AC O99071;
 DT 01-MAY-1999 (TReMBLrel. 10, Created)
 DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
 DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
 DE Cytochrome b (Fragment).
 GN Name=cytb;
 OS Pronothobranchius kiyawensis.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha; OC Cyprinodontiformes; Aplocheilidae; Aplocheilinae; Pronothobranchius. OX NCBI_TaxID=60306;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Aquarium strain;
 RX MEDLINE=99214037; PubMed=10196077; DOI=10.1006/mpev.1998.0566;
 RA Murphy W.J., Collier G.E.;
 RT "Phylogenetic relationships of African killifishes in the genera Aphyosemon and Fundulopanchax inferred from mitochondrial DNA sequences.";
 RL Mol. Phylogenet. Evol. 11:351-360(1999).
 CC -!- FUNCTION: Component of the ubiquinol-cytochrome c reductase complex (complex III or cytochrome b-c1 complex), which is a respiratory chain that generates an electrochemical potential coupled to ATP synthesis (By similarity).
 CC -!- COPACTOR: Binds two heme groups non-covalently. Heme 1 (or BL or b562) is low-potential and absorbs at about 562 nm, and heme 2 (or BH or b566) is high-potential and absorbs at about 566 nm (By similarity).
 CC -!- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b, cytochrome c1 and the Rieske protein (By similarity).
 CC -!- SIMILARITY: Belongs to the cytochrome b family.
 DR EMBL; AF002285; AAD00946.1; -.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO; GO:0005739; C:mitochondrion; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR005797; Cytb_b6_N.
 DR Pfam; PF00033; Cytochrom_B_N; 1.
 DR PROSITE; PS00192; CYTOCHROME_B_HEME; 1.
 KW Electron transport; Heme; Mitochondrion; Respiratory chain; Transmembrane; Transport.
 FT NON_TER 1
 FT NON_TER 120 120
 SQ SEQUENCE 120 AA; 13591 MW; 54D1FD7D50C1538CE CRC64;
 Query Match 100.0%; Score 25; DB 2; Length 120;
 Best Local Similarity 100.0%; Pred. No. 98;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR005797; Cytb_b6_N.
 DR Pfam; PF00033; Cytochrom_B_N; 1.
 DR PROSITE; PS00192; CYTOCHROME_B_HEME; 1.
 KW Electron transport; Heme; Mitochondrion; Respiratory chain; Transmembrane; Transport.
 FT NON_TER 1
 FT NON_TER 120 120
 SQ SEQUENCE 120 AA; 13600 MW; EECADA4FE9B9639B CRC64;
 Query Match 100.0%; Score 25; DB 2; Length 120;
 Best Local Similarity 100.0%; Pred. No. 98;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SSVMH 5
 Db 60 SSVMH 64
 RESULT 10
 Q99071 PRELIMINARY; PRT; 120 AA.
 AC O99071;
 DT 01-MAY-1999 (TReMBLrel. 10, Created)
 DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
 DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
 DE Cytochrome b (Fragment).
 GN Name=cytb;
 OS Pronothobranchius kiyawensis.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha; OC Cyprinodontiformes; Aplocheilidae; Aplocheilinae; Pronothobranchius. OX NCBI_TaxID=60306;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Aquarium strain;
 RX MEDLINE=99214037; PubMed=10196077; DOI=10.1006/mpev.1998.0566;
 RA Murphy W.J., Collier G.E.;
 RT "Phylogenetic relationships of African killifishes in the genera Aphyosemon and Fundulopanchax inferred from mitochondrial DNA sequences.";
 RL Mol. Phylogenet. Evol. 11:351-360(1999).
 CC -!- FUNCTION: Component of the ubiquinol-cytochrome c reductase complex (complex III or cytochrome b-c1 complex), which is a respiratory chain that generates an electrochemical potential coupled to ATP synthesis (By similarity).
 CC -!- COPACTOR: Binds two heme groups non-covalently. Heme 1 (or BL or b562) is low-potential and absorbs at about 562 nm, and heme 2 (or BH or b566) is high-potential and absorbs at about 566 nm (By similarity).
 CC -!- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b, cytochrome c1 and the Rieske protein (By similarity).
 CC -!- SIMILARITY: Belongs to the cytochrome b family.
 DR EMBL; AF002285; AAD00946.1; -.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO; GO:0005739; C:mitochondrion; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR005797; Cytb_b6_N.
 DR Pfam; PF00033; Cytochrom_B_N; 1.
 DR PROSITE; PS00192; CYTOCHROME_B_HEME; 1.
 KW Electron transport; Heme; Mitochondrion; Respiratory chain; Transmembrane; Transport.
 FT NON_TER 1
 FT NON_TER 120 120
 SQ SEQUENCE 120 AA; 13591 MW; 54D1FD7D50C1538CE CRC64;
 Query Match 100.0%; Score 25; DB 2; Length 120;
 Best Local Similarity 100.0%; Pred. No. 98;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
 Db 60 SSVMH 64

RESULT 11

ID Q64GF1 PRELIMINARY; PRT; 122 AA.
 AC Q64GF1; 28, Created)
 DT 25-OCT-2004 (TREMELREL. 28, Last sequence update)
 DT 25-OCT-2004 (TREMELREL. 28, Last sequence update)
 DE Cytochrome b (Fragment).
 GN Names=cytb;
 OS Plethodon elongatus (Del Norte salamander).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Caudata; Salamandroidae; Plethodontidae;
 OC Plethodontinae; Plethodontini; Plethodon.
 OX NCBI_TaxID=46761;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RAF0245;
 RA Mead L.S., Clayton D.R., Nauman R.S., Olson D.H., Pfrender M.E.;
 RT "Newly discovered populations of salamanders from Siskiyou County
 California represent a species distinct from Plethodon stormi.";
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY68284; AAU21592.1; --
 KW Mitochondrion.
 FT NON_TER 1
 FT NON_TER 122
 SQ SEQUENCE 122 AA; 14064 MW; 841DB00A9785671D CRC64;

Query Match 100.0%; Score 25; DB 2; Length 122;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
 Db 66 SSVMH 70

RESULT 12

ID Q8MLH7 PRELIMINARY; PRT; 128 AA.
 AC Q8MLH7; 22, Created)
 DT 01-OCT-2002 (TREMELREL. 22, Last sequence update)
 DT 01-OCT-2002 (TREMELREL. 22, Last sequence update)
 DT 01-NAR-2004 (TREMELREL. 26, Last annotation update)
 DE Cytochrome b (Fragment).
 OS Mesoplon peruvianus (Peruvian beaked whale).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Cetacea; Odontoceti; Ziphiidae;
 OC Mesoplon.
 OX NCBI_TaxID=27617;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA van Helden A.L., Baker A.N., Dalebout M.L., Reyes J.C.,
 RA Van Waerebeek K., Baker C.S.;
 RT "Resurrection of Mesoplon traversi (Gray, 1874), senior synonym of
 M. bahamondi Reyes, Van Waerebeek, Cardenas and Yanez, 1995 (Cetacea:
 Ziphiidae).";
 RL Mar. Mamm. Sci. 18:609-621(2002).
 CC -!- FUNCTION: Component of the ubiquinol-cytochrome c reductase
 complex (complex III or cytochrome b-c1 complex), which is a
 respiratory chain that generates an electrochemical potential
 coupled to ATP synthesis (By similarity).
 CC -!- COFACTOR: Binds two heme groups non-covalently. Heme 1 (or BL or
 b562) is low-potential and absorbs at about 562 nm, and heme 2 (or
 BH or b566) is high-potential and absorbs at about 566 nm (By
 similarity).

similarity).
 -!- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
 cytochrome c1 and the Rieske protein (By similarity).
 -!- SIMILARITY: Belongs to the cytochrome b family.
 DR EMBL; AF492414; AAM74073.1; --
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO; GO:0005739; C:mitochondrion; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR005797; Cytb_b6_N.
 DR Pfam; PF00033; Cytochrom B_N; 1.
 DR PROSITE; PS00192; CYTOCHROME B_HEME; 1.
 KW Electron transport; Heme; Mitochondrion; Respiratory chain;
 FT Transmembrane; Transport.
 FT NON_TER 128
 SQ SEQUENCE 128 AA; 14611 MW; 5BF9B08C2A6CC991 CRC64;

Query Match 100.0%; Score 25; DB 2; Length 128;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5
 Db 64 SSVMH 68

RESULT 13

ID Q6Y4C8 PRELIMINARY; PRT; 128 AA.
 AC Q6Y4C8;
 DT 05-JUL-2004 (TREMELREL. 27, Created)
 DT 05-JUL-2004 (TREMELREL. 27, Last sequence update)
 DT 05-JUL-2004 (TREMELREL. 27, Last annotation update)
 DE Cytochrome b (Fragment).
 GN Name=cytb;
 OS Plethodon elongatus (Del Norte salamander).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Caudata; Salamandroidae; Plethodontidae;
 OC Plethodontinae; Plethodontini; Plethodon.
 OX NCBI_TaxID=46761;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed=14653796; DOI=10.1046/j.1365-294X.2003.02030.x;
 RA Mahoney M.J.;
 RT "Molecular systematics and phylogeography of the Plethodon elongatus
 species group: combining phylogenetic and population genetic methods
 to investigate species history.";
 RL Mol. Ecol. 13:149-166(2004).
 CC -!- FUNCTION: Component of the ubiquinol-cytochrome c reductase
 complex (complex III or cytochrome b-c1 complex), which is a
 respiratory chain that generates an electrochemical potential
 coupled to ATP synthesis (By similarity).
 CC -!- COFACTOR: Binds two heme groups non-covalently. Heme 1 (or BL or
 b562) is low-potential and absorbs at about 562 nm, and heme 2 (or
 BH or b566) is high-potential and absorbs at about 566 nm (By
 similarity).

-!- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
 cytochrome c1 and the Rieske protein (By similarity).
 -!- SIMILARITY: Belongs to the cytochrome b family.
 DR EMBL; AY183769; AAO63634.1; --
 DR EMBL; AY183769; AAO63633.1; --
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO; GO:0005739; C:mitochondrion; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR005797; Cytb_b6_N.
 DR Pfam; PF00033; Cytochrom B_N; 1.
 DR PROSITE; PS00192; CYTOCHROME B_HEME; 1.

KW Electron transport; Heme; Mitochondrion; Respiratory chain;
KW Transmembrane; Transport.

FT NON_TER 1

FT NON_TER 128

SQ SEQUENCE 128 AA; 14630 MW; F5485853845ADC2B CRC64;

Query Match 100.0%; Score 25; DB 2; Length 128;

Best Local Similarity 100.0%; Pred. No. 1.1e+02; Indels 0; Gaps 0;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5

Db 58 SSVMH 62

RESULT 14

ID Q6Y4D1

AC Q6Y4D1 PRELIMINARY; PRT; 128 AA.

DT 05-JUL-2004 (TRENBLrel. 27, Created)

DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)

DE 05-JUL-2004 (TRENBLrel. 27, Last annotation update)

DE Cytochrome b (Fragment).

GN Name=cytb;

OS Plethodon elongatus (Del Norte salamander).

OG Mitochondrion.

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Caudata; Salamandroidae; Plethodontidae;

OC Plethodontinae; Plethodontini; Plethodon.

OX NCBI_TaxID=46761;

RN [1]

RP SEQUENCE FROM N.A.

RX PubMed=14653796; DOI=10.1046/j.1365-294X.2003.02030.x;

RA Mahoney M.J.;

RT "Molecular systematics and phylogeography of the Plethodon elongatus

species group: combining phylogenetic and population genetic methods

to investigate species history.";

RL Mol. Ecol. 13:149-166(2004).

CC -!- FUNCTION: Component of the ubiquinol-cytochrome c reductase

complex (complex III or cytochrome b-c1 complex), which is a

respiratory chain that generates an electrochemical potential

coupled to ATP synthesis (By similarity).

CC -!- COFACTOR: Binds two heme groups non-covalently. Heme 1 (or BL or

b562) is low-potential and absorbs at about 562 nm, and heme 2 (or

BH or b566) is high-potential and absorbs at about 566 nm (By

similarity).

CC -!- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,

cytochrome c1 and the Rieske protein (By similarity).

CC -!- SIMILARITY: Belongs to the cytochrome b family.

DR EMBL; AY183766; AAO63631.1; -.

DR EMBL; AY183765; AAO63630.1; -.

DR GO; GO:0016021; C:integral to membrane; IEA.

DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.

DR GO; GO:0005739; C:mitochondrion; IEA.

DR GO; GO:0016491; F:oxidoreductase activity; IEA.

DR GO; GO:0006118; P:electron transport; IEA.

DR GO; GO:0006810; P:transport; IEA.

DR InterPro; IPR005797; Cytb_b6_N.

DR Pfam; PF00033; Cytochrom B_N; 1.

DR PROSITE; PS00192; CYTOCHROME B HEME; 1.

KW Electron transport; Heme; Mitochondrion; Respiratory chain;

KW Transmembrane; Transport.

FT NON_TER 1

FT NON_TER 128

SQ SEQUENCE 128 AA; 14664 MW; 77CA5853845AD225 CRC64;

Query Match

Best Local Similarity 100.0%; Score 25; DB 2; Length 128;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5

Db 58 SSVMH 62

RESULT 15

Q9G155

AC Q9G155 PRELIMINARY; PRT; 128 AA.

DT 01-MAR-2001 (TRENBLrel. 16, Created)

DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)

DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)

DE Cytochrome b (Fragment).

OS Entosphenus minimus.

OG Mitochondrion.

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;

OC Petromyzontiformes; Petromyzontidae; Entosphenus.

OX NCBI_TaxID=126726;

RN [1]

RP SEQUENCE FROM N.A.

RA Lorion C.M.; Markle D.F.; Reid S.B.; Docker M.F.;

RT "Redescription of the presumed-extinct Miller Lake lamprey, Lampetra

minima.";

RL Copeia 100:1019-1028(2000).

CC -!- FUNCTION: Component of the ubiquinol-cytochrome c reductase

complex (complex III or cytochrome b-c1 complex), which is a

respiratory chain that generates an electrochemical potential

coupled to ATP synthesis (By similarity).

CC -!- COFACTOR: Binds two heme groups non-covalently. Heme 1 (or BL or

b562) is low-potential and absorbs at about 562 nm, and heme 2 (or

BH or b566) is high-potential and absorbs at about 566 nm (By

similarity).

CC -!- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,

cytochrome c1 and the Rieske protein (By similarity).

CC -!- SIMILARITY: Belongs to the cytochrome b family.

DR EMBL; AF257132; AAG13282.1; -.

DR EMBL; AF257130; AAG13280.1; -.

DR GO; GO:0016021; C:integral to membrane; IEA.

DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.

DR GO; GO:0005739; C:mitochondrion; IEA.

DR GO; GO:0016491; F:oxidoreductase activity; IEA.

DR GO; GO:0006118; P:electron transport; IEA.

DR GO; GO:0006810; P:transport; IEA.

DR InterPro; IPR005797; Cytb_b6_N.

DR Pfam; PF00033; Cytochrom B_N; 1.

DR PROSITE; PS00192; CYTOCHROME B HEME; UNKNOWN 1.

KW Electron transport; Heme; Mitochondrion; Respiratory chain;

KW Transmembrane; Transport.

FT NON_TER 128

FT NON_TER 128

SQ SEQUENCE 128 AA; 14301 MW; BABCB4B81EEC4C4A CRC64;

Query Match

Best Local Similarity 100.0%; Score 25; DB 2; Length 128;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVMH 5

Db 68 SSVMH 72

Search completed: April 28, 2005, 18:24:23

Job time : 35.3438 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 01:32:26 ; Search time 109.377 Seconds
(without alignments)
43.915 Million cell updates/sec

Title: US-10-088-639A-2_COPY_177_193
Perfect score: 85
Sequence: 1 VISENGRTINADVSKG 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database : A_Geneseq_25Jan04: *
1: Geneseqp1980s: *
2: Geneseqp1990s: *
3: Geneseqp2000s: *
4: Geneseqp2001s: *
5: Geneseqp2002s: *
6: Geneseqp2003as: *
7: Geneseqp2003bs: *
8: Geneseqp2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	85	100.0	249	4	AAB68087 An anti-a
2	67	78.8	120	2	AAR47042 Sequence
3	65	76.5	17	5	AU82607 Llama CDR
4	63	74.1	17	6	ADA91299 MS-R Fab/
5	63	74.1	17	6	ADA90004 Anti-Abet
6	63	74.1	17	6	ADA91201 MS-R Fab/
7	63	74.1	17	6	ADA90434 MS-R Roche
8	63	74.1	17	6	ADA91199 MS-R Fab/
9	63	74.1	17	6	ADA911301 MS-R Fab/
10	63	74.1	120	6	ADA90190 Anti-Abet
11	63	74.1	120	6	ADA90188 Anti-Abet
12	62	72.9	17	6	ADA90010 Anti-Abet
13	62	72.9	17	6	ADA91195 MS-R Fab/
14	62	72.9	17	6	ADA91198 MS-R Fab/
15	62	72.9	17	6	ADA90000 Anti-Abet
16	62	72.9	17	6	ADA91193 MS-R Fab/
17	62	72.9	17	6	ADA90430 MS-R Roche
18	62	72.9	17	6	ADA91196 MS-R Fab/
19	62	72.9	17	6	ADA90440 MS-R Roche
20	62	72.9	120	6	ADA90186 Anti-Abet
21	62	72.9	120	6	ADA90184 Anti-Abet
22	61	71.8	17	6	ADA90021 Anti-Abet
23	61	71.8	17	6	ADA90580 MS-R Roche
24	61	71.8	121	4	AAG62968 Amino aci
25	61	71.8	248	5	ABP45637 Human Bly

26	60	70.6	17	6	ADA90581 MS-R Roche
27	60	70.6	17	6	ADA90022 Anti-Abet
28	60	70.6	115	7	ADD28077 Lymphoma
29	60	70.6	120	2	AAR30765 Consensus
30	59	69.4	17	6	ADA90439 MS-R Roche
31	59	69.4	17	6	ADA90009 Anti-Abet
32	59	69.4	122	4	AAU02547 Anti-adiop
33	59	69.4	129	4	AAG93631 Human ant
34	59	69.4	129	6	ABO27438 Anti-Rh(D
35	58	68.2	17	6	ADA90425 MS-R Roche
36	58	68.2	17	6	ADA91069 MS-R Fab/
37	58	68.2	17	6	ADA89995 Anti-Abet
38	58	68.2	17	6	ADA90582 MS-R Roche
39	58	68.2	17	6	ADA91304 MS-R Fab/
40	58	68.2	17	6	ADA91067 MS-R Fab/
41	58	68.2	17	6	ADA91187 MS-R Fab/
42	58	68.2	17	6	ADA91302 MS-R Fab/
43	58	68.2	17	6	ADA90023 Anti-Abet
44	58	68.2	17	6	ADA91189 MS-R Fab/
45	58	68.2	120	6	ADA90180 Anti-Abet
46	58	68.2	120	6	ADA90192 Anti-Abet
47	58	68.2	120	6	ADA89918 Anti-Abet
48	58	68.2	240	4	AAB46021 Human MUC
49	58	68.2	240	4	AAB46039 Human TF
50	58	68.2	250	4	AAU00212 Monoclonal
51	58	68.2	261	5	ABP45492 Human Bly
52	57	67.1	17	6	ADA90583 MS-R Roche
53	57	67.1	17	6	ADA90024 Anti-Abet
54	57	67.1	22	2	AAU06694 Ab2 varia
55	57	67.1	40	6	ABJ25337 SAC isola
56	57	67.1	98	2	AAW59614 Anti-RSV
57	57	67.1	98	3	AAB40075 Anti-hiL1
58	57	67.1	98	3	AAB40113 Anti-hiL1
59	57	67.1	98	4	AAB48022 Heavy cha
60	57	67.1	98	5	ABG78188 Human Fv
61	57	67.1	98	5	ABG78204 Human Fv
62	57	67.1	98	5	ABG77149 Germline
63	57	67.1	98	5	ABG91879 Human ant
64	57	67.1	98	5	ABG91895 Human ant
65	57	67.1	98	6	ABO27084 Human ger
66	57	67.1	98	7	ADD28042 Lymphoma
67	57	67.1	98	7	ADD28076 Lymphoma
68	57	67.1	109	7	ADC60996 Human ant
69	57	67.1	114	7	ADD28045 Lymphoma
70	57	67.1	117	2	AAR66305 Human imm
71	57	67.1	117	5	AAO15187 Dig3 anti
72	57	67.1	120	4	AAU02559 Anti-adiop
73	57	67.1	122	6	ABR55779 Heavy cha
74	57	67.1	122	6	ABR55821 Heavy cha
75	57	67.1	122	6	ABR55775 Heavy cha
76	57	67.1	123	2	AAW08582 Human ant
77	57	67.1	124	5	AAO18439 Anti-GD2
78	57	67.1	124	5	AAO18440 Anti-GD2
79	57	67.1	131	4	AAG65566 Amino aci
80	57	67.1	131	5	AAO15843 Human imm
81	57	67.1	196	3	AAV56862 Human imm
82	57	67.1	236	4	AB46055 Human TF
83	57	67.1	236	5	ABP45999 Human Bly
84	57	67.1	237	5	ABP46101 Human Bly
85	57	67.1	237	5	ABP46017 Human Bly
86	57	67.1	237	5	ABP46032 Human Bly
87	57	67.1	237	5	ABP46107 Human Bly
88	57	67.1	237	5	ABP45995 Human Bly
89	57	67.1	237	5	ABP46093 Human Bly
90	57	67.1	237	5	ABP46100 Human Bly
91	57	67.1	237	5	ABP45992 Human Bly
92	57	67.1	237	5	ABP46008 Human Bly
93	57	67.1	237	5	ABP46025 Human Bly
94	57	67.1	237	5	ABP46103 Human Bly
95	57	67.1	237	5	ABP46104 Human Bly
96	57	67.1	237	5	ABP46016 Human Bly
97	57	67.1	237	5	ABP45994 Human Bly
98	57	67.1	237	5	ABP46006 Human Bly

99 57 67.1 237 5 ABP46009 Human Bly
 100 57 67.1 237 5 ABP46029 Human Bly
 101 57 67.1 237 5 ABP46099 Human Bly
 102 57 67.1 238 5 ABP46013 Human Bly
 103 57 67.1 239 5 ABP46023 Human Bly
 104 57 67.1 240 4 AAB46043 Human TF
 105 57 67.1 240 5 ABP46005 Human Bly
 106 57 67.1 240 5 ABP46033 Human Bly
 107 57 67.1 240 5 ABP46094 Human Bly
 108 57 67.1 240 5 ABP46037 Human Bly
 109 57 67.1 240 5 ABP46014 Human Bly
 110 57 67.1 240 5 ABP46019 Human Bly
 111 57 67.1 240 5 ABP46097 Human Bly
 112 57 67.1 240 5 ABP46018 Human Bly
 113 57 67.1 240 5 ABP45996 Human Bly
 114 57 67.1 240 5 ABP46034 Human Bly
 115 57 67.1 240 5 ABP46106 Human Bly
 116 57 67.1 240 5 ABP46030 Human Bly
 117 57 67.1 240 5 ABP46102 Human Bly
 118 57 67.1 245 2 AAY06714 Antibody
 119 57 67.1 248 5 ABP45963 Human Bly
 120 57 67.1 248 5 ABP45954 Human Bly
 121 57 67.1 250 5 ABP45168 Human Bly
 122 57 67.1 473 5 ABG77162 Germine
 123 56 65.9 17 2 AAY05053 Tumour an
 124 56 65.9 17 6 ADA90423 MS-Roche
 125 56 65.9 17 6 ADA90438 MS-Roche
 126 56 65.9 17 6 ADA89993 Anti-Abet
 127 56 65.9 17 6 ADA90008 Anti-Abet
 128 56 65.9 17 7 ADC82824 CDR regio
 129 56 65.9 40 6 ABJ25311 SAC isola
 130 56 65.9 98 7 ADD28072 Lymphoma
 131 56 65.9 119 5 ABG32513 Human VH
 132 56 65.9 124 4 AAG93561 Human ant
 133 56 65.9 124 6 ABR54879 IgG kappa
 134 56 65.9 124 6 ABR54878 IgG kappa
 135 56 65.9 124 6 ABO27368 Anti-Rh(D
 136 56 65.9 240 4 AAB46002 Human MOC
 137 56 65.9 240 4 AAB46001 Human MOC
 138 56 65.9 254 5 ABP45785 Human Bly
 139 55 64.7 17 4 AAB61288 Anti-TANG
 140 55 64.7 17 6 ABUL1254 Human TAN
 141 55 64.7 17 6 ADA90427 MS-Roche
 142 55 64.7 17 6 ADA90012 Anti-Abet
 143 55 64.7 17 6 ADA89997 Anti-Abet
 144 55 64.7 17 6 ADA90442 MS-Roche
 145 55 64.7 40 6 ABJ25321 SAC isola
 146 55 64.7 40 6 ABJ25332 SAC isola
 147 55 64.7 107 7 ADD28087 Lymphoma
 148 55 64.7 121 2 AAR13182 Variable
 149 55 64.7 121 2 AAR88848 Human ant
 150 55 64.7 123 6 ABR57130 Camel ant

ALIGNMENTS

RESULT 1
 AAB68087
 ID AAB68087 standard; protein; 249 AA.

XX AC AAB68087;

XX DT 09-JUL-2001 (first entry)

XX DE An anti-alpha6beta4 integrin light chain linked to a heavy chain.

XX KW Gastrointestinal epithelial tumour cell; alpha6beta4 integrin;

XX KW tumour-associated antigen; metastatic disease; malignant disease;

XX KW tumour typing; tumour screening; tumour.

XX OS Synthetic.

OS Macaca fascicularis.

XX FH Location/Qualifiers
 FT 23. .33
 FT /note= "Complementarity determining region (CDR) 1 of the
 light chain"
 FT 49. .55
 FT /note= "Complementarity determining region (CDR) 2 of the
 light chain"
 FT 88. .98
 FT /note= "Complementarity determining region (CDR) 3 of the
 light chain"
 FT 110. .127
 FT /note= "linker"
 FT 158. .162
 FT /note= "Complementarity determining region (CDR) 1 of the
 heavy chain"
 FT 177. .193
 FT /note= "Complementarity determining region (CDR) 1 of the
 heavy chain"
 FT 226. .238
 FT /note= "Complementarity determining region (CDR) 1 of the
 heavy chain"
 WO200130854-A2.
 03-MAY-2001.
 26-OCT-2000; 2000WO-SE002082.
 28-OCT-1999; 99SE-00003895.
 (ACTI-) ACTIVE BIOTECH AB.
 Brodin TN, Karlstroem PJ, Ohlsson LG, Tordsson MJ, Kearney PP;
 Nilsson BHK;
 WPI: 2001-308619/32.
 N-PSDB; AAF84797.

Novel antibody for diagnosis, treatment of human metastatic and malignant
 diseases, has binding structure for target structure displayed on cell
 surface of human gastrointestinal epithelial tumor and normal cells.

Claim 1; Page 55-56; 75pp; English.

The present sequence represents a Monkey antibody light chain linked to a
 heavy chain. The antibody binds to a target structure displayed in and on
 the cell surface of human gastrointestinal epithelial tumour cells and in
 a subpopulation of normal human gastrointestinal epithelial cells. The
 target structure especially comprises alpha6beta4 integrin. This is a
 tumour-associated antigen. The antibody, and its fragments, are useful
 for treating conditions based on anti-angiogenic mechanism in humans.
 They are useful for treating human metastatic and malignant disease,
 in vitro, in vivo diagnosis and prognosis of human malignant disease,
 comprising tumour typing, tumour screening, tumour diagnosis and
 prognosis and monitoring premalignant conditions. Quantitative in vivo
 diagnosis is carried out by determining the localization of antibody to
 tumour deposits in humans

Sequence 249 AA;

Query Match 100.0%; Score 85; DB 4; Length 249;

Best Local Similarity 100.0%; Pred. No. 2e-06;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 VISENGRTINYADSVKG 17

Db 177 VISENGRTINYADSVKG 193
 |||||

RESULT 2

AAR47042

ID AAR47042 standard; peptide; 120 AA.

XX AC AAR47042;
 XX DT 25-MAR-2003 (revised)
 XX DT 02-SEP-1994 (first entry)
 XX DE Sequence of the consensus antibody variable domain of the heavy chain.
 XX KW Monoclonal antibody; heavy chain; variable domain; consensus.
 XX OS Synthetic.
 XX PN WO9404679-A1.
 XX PD 03-MAR-1994.
 XX PF 20-AUG-1993; 93WO-US007832.
 XX PR 21-AUG-1992; 92US-00934373.
 XX PA (GETH) GENENTECH INC.
 XX PI Carter PJ, Presta LG;
 XX DR WPI; 1994-083196/10.
 XX CC Preparation of improved humanised antibodies - by comparison of consensus
 PT and import complementarity determining regions and framework region
 PT sequences, e.g. to humanise murine.
 XX PS Claim 16; Page 108; 126pp; English.
 XX CC The consensus variable domain sequences are derived from the most
 CC abundant subclasses in the sequence compilation of Kabat et al., Sequences
 CC of Proteins of Immunological Interest, National Institutes of Health,
 CC Bethesda MD (1987), namely, VL-kappa subgroup I and VH group III. In such
 CC IGG-gamma-1 human consensus sequences, the VI consensus domain has the AA
 CC sequence in AAR47041, and the VH consensus domain has the AA sequence in
 CC AAR47042. (Updated on 25-MAR-2003 to correct PN field.)
 XX SQ Sequence 120 AA;
 Query Match 78.8%; Score 67; DB 2; Length 120;
 Best Local Similarity 82.4%; Pred. No. 0.0012;
 Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 VISENGRTINVADSVKG 17
 DB 50 VISENGGTRYADSVKG 66
 RESULT 3
 AAU82607
 ID AAU82607 standard; peptide; 17 AA.
 AC AAU82607;
 XX 29-AUG-2003 (revised)
 DT 23-APR-2002 (first entry)
 XX DE Llama CDR2 region variable heavy chain fragment #61.
 XX KW Llama; phage display library; variable heavy domain fragment; VHH; VH;
 KW sdAb fragment; single domain anti-idiotypic antibody fragment;
 KW phage display technology; immune system response; CDR1/H1; CDR2; CDR3;
 KW complementarity determining region.
 XX OS Lama glama.
 XX PN WO200190190-A2.
 XX PD 29-NOV-2001.
 XX

PF 25-MAY-2001; 2001WO-CA000763.
 XX 26-MAY-2000; 2000US-0207234P.
 XX (CANA) NAT RES COUNCIL CANADA.
 XX PI Tanha J, Dubuc G, Narang S;
 XX DR WPI; 2002-083093/11.
 XX PT New phage display library of variable heavy domain antigen-binding
 PT fragments derived from llama antibodies, useful for in vitro selection
 PT against any antigen of interest as a target.
 XX PS Disclosure; Page 19a; 46pp; English.
 XX CC The present invention relates to a phage display library of variable
 CC heavy domain (VHH or VH) fragments (sdAb fragments) derived from llama
 CC antibodies. The library is useful for in vitro selection against any
 CC antigen of interest as a target. Single domain anti-idiotypic antibody
 CC fragments are isolated from the library using phage display technology
 CC and an antibody serving as an antigen. Such anti-idiotypic antibody
 CC fragments have great potential in evoking the immune system response to
 CC pathological antigens and in vaccine development. The large size of the
 CC library considerably increases the probability of isolating from it
 CC antigen-binding fragments having high affinity to almost any
 CC predetermined target (antigen of interest). The library eliminates the
 CC development of anti-idiotypic antibodies by immunisation and allows
 CC isolation of anti-idiotypic antibodies. AAU82435-AAU82635 represent the
 CC llama heavy chain domain fragments of the invention. (Updated on 29-AUG-
 CC 2003 to standardise OS field)
 XX SQ Sequence 17 AA;
 Query Match 76.5%; Score 65; DB 5; Length 17;
 Best Local Similarity 70.6%; Pred. No. 0.00025;
 Matches 12; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 1 VISENGRTINVADSVKG 17
 DB 1 LLSRSGRTTYADSVKG 17
 RESULT 4
 ADA91299
 ID ADA91299 standard; peptide; 17 AA.
 XX ADA91299;
 AC ADA91299;
 XX 20-NOV-2003 (first entry)
 DT MS-R Fab/antibody related peptide #347.
 DE antibody molecule; antibody; beta-A4 peptide; Abeta4; neuroprotective;
 XX nootropic; antiparkinsonian; gene therapy; amyloidogenesis;
 KW amyloid-plaque formation; beta-amyloid plaque; immunisation; dementia;
 KW Alzheimer's disease; motor neuropathy; Down's syndrome;
 KW Creutzfeldt Jacob disease; hereditary cerebral haemorrhage; amyloidosis;
 KW Parkinson's disease; HIV-related dementia; amyotrophic lateral sclerosis;
 KW neuronal disorder; aging.
 XX OS Synthetic.
 OS Homo sapiens.
 XX WO2003070760-A2.
 PN 28-AUG-2003.
 XX 20-FEB-2003; 2003WO-EP001759.
 PF 20-FEB-2002; 2002EP-00003844.
 XX (HOFF) HOFFMANN LA ROCHE & CO AG F.
 XX PA

PA (MORP-) MORPHOSYS AG.
XX
PI Bardroff M, Bohrmann B, Brockhaus M, Huber W, Kretzschmar T;
PI Loehning C, Loetscher H, Nordstedt C, Rothe C;
XX
XX WPI; 2003-663848/62.
XX
XX New antibody molecule capable of specifically recognizing two regions of
PT the beta-A4 peptide, useful for diagnosing, preventing or treating
PT diseases associated with amyloidogenesis or amyloid-plaque formation
PT (e.g. dementia).
XX
XX Example 13; Page 93; 312pp; English.
XX
XX The present invention describes an antibody molecule (I) capable of
CC specifically recognising two regions of the beta-A4 peptide/Abeta4. The
CC first region comprises the amino acid sequence Ala-Glu-Phe-Arg-His-Asp-
CC Ser-Gly-Tyr ADA89886 or its fragment, and the second region comprises the
CC amino acid sequence Val-His-His-Gln-Lys-Leu-Val-Phe-Phe-Ala-Glu-Asp-Val-
CC Gly ADA89887 or its fragment. Also described: (1) a nucleic acid molecule
CC encoding (1); (2) a vector comprising the nucleic acid of (1); (3) a host
CC cell comprising the vector of (2); (4) preparing (1), comprising
CC culturing the host cell of (3) under conditions that allow synthesis of
CC (1) and recovering (1) from the culture; (5) a composition comprising (1)
CC or an antibody molecule produced by method (4); (6) a kit comprising (1),
CC (1); (8) testing the resulting Fab optimisation library by panning
CC against Abeta/Abeta4; (9) identifying optimised clones; (10) expressing
CC of selected, optimised clones; (11) preparing a pharmaceutical
CC optimised antibody/antibody molecule with a carrier; and (12) a
CC pharmaceutical composition prepared by method (8). (1) has
CC neuroprotective, neurotropic and antiparkinsonian activities, and can be
CC used in gene therapy. The antibody molecule (I), nucleic acid molecule,
CC vector or host is useful in preparing a pharmaceutical composition for
CC the prevention and/or treatment of a disease associated with
CC amyloidogenesis and/or amyloid-plaque formation. The antibody molecule
CC may also be used in preparing a diagnostic composition for the detection
CC of the disease mentioned above. The antibody is used for the
CC disintegration of beta-amyloid plaques or for passive immunisation
CC against beta-amyloid plaque formation. In particular, the disease is
CC dementia, Alzheimer's disease, motor neuropathy, Down's syndrome,
CC Creutzfeldt Jacob disease, hereditary cerebral haemorrhage with
CC amyloidosis Dutch type, Parkinson's disease, HIV-related dementia,
CC amyotrophic lateral sclerosis or neuronal disorders related to aging. The
CC present sequence is used in the exemplification of the present invention.
XX
XX Sequence 17 AA;
XX
XX Query Match 74.1%; Score 63; DB 6; Length 17;
XX Best Local Similarity 76.5%; Pred. No. 0.00056;
XX Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
XX
XX 1 VISENGRTINYADSVKG 17
XX ||||| : |||||
XX 1 VISETGKNIYADSVKG 17
XX
XX RESULT 5
XX ID ADA90004
XX ID ADA90004 standard; peptide; 17 AA.
XX
XX ADA90004;
XX
XX 20-NOV-2003 (first entry)
XX
XX Anti-Abeta antibody related amino acid sequence SEQ ID NO:119.
XX
XX antibody molecule; antibody; beta-A4 peptide; Abeta4; neuroprotective;
XX neurotropic; antiparkinsonian; gene therapy; amyloidogenesis;
XX amyloid-plaque formation; beta-amyloid plaque; immunisation; dementia;
XX Alzheimer's disease; motor neuropathy; Down's syndrome;
XX Creutzfeldt Jacob disease; hereditary cerebral haemorrhage; amyloidosis;

KW Parkinson's disease; HIV-related dementia; amyotrophic lateral sclerosis;
KW neuronal disorder; aging.
XX
XX Synthetic.
OS Homo sapiens.
XX
XX WO2003070760-A2.
XX
XX 28-AUG-2003.
XX
XX 20-FEB-2003; 2003WO-EP001759.
XX
XX 20-FEB-2002; 2002EP-00003844.
XX
XX (HOFF) MORPHOSYS AG.
PA
PA (MORP-) MORPHOSYS AG.
XX
XX Bardroff M, Bohrmann B, Brockhaus M, Huber W, Kretzschmar T;
XX Loehning C, Loetscher H, Nordstedt C, Rothe C;
XX
XX WPI; 2003-663848/62.
XX
XX New antibody molecule capable of specifically recognizing two regions of
PT the beta-A4 peptide, useful for diagnosing, preventing or treating
PT diseases associated with amyloidogenesis or amyloid-plaque formation
PT (e.g. dementia).
XX
XX Disclosure; Page 189-190; 312pp; English.
XX
XX The present invention describes an antibody molecule (I) capable of
CC specifically recognising two regions of the beta-A4 peptide/Abeta4. The
CC first region comprises the amino acid sequence Ala-Glu-Phe-Arg-His-Asp-
CC Ser-Gly-Tyr ADA89886 or its fragment, and the second region comprises the
CC amino acid sequence Val-His-His-Gln-Lys-Leu-Val-Phe-Phe-Ala-Glu-Asp-Val-
CC Gly ADA89887 or its fragment. Also described: (1) a nucleic acid molecule
CC encoding (1); (2) a vector comprising the nucleic acid of (1); (3) a host
CC cell comprising the vector of (2); (4) preparing (1), comprising
CC culturing the host cell of (3) under conditions that allow synthesis of
CC (1) and recovering (1) from the culture; (5) a composition comprising (1)
CC or an antibody molecule produced by method (4); (6) a kit comprising (1),
CC nucleic acid of (1), vector of (2) or host cell of (3); (7) optimising
CC (1); (8) testing the resulting Fab optimisation library by panning
CC against Abeta/Abeta4; (9) identifying optimised clones; (10) expressing
CC of selected, optimised clones; (11) preparing a pharmaceutical
CC composition, comprising optimisation of (1), and formulating the
CC optimised antibody/antibody molecule with a carrier; and (12) a
CC pharmaceutical composition prepared by method (8). (1) has
CC neuroprotective, neurotropic and antiparkinsonian activities, and can be
CC used in gene therapy. The antibody molecule (I), nucleic acid molecule,
CC vector or host is useful in preparing a pharmaceutical composition for
CC the prevention and/or treatment of a disease associated with
CC amyloidogenesis and/or amyloid-plaque formation. The antibody molecule
CC may also be used in preparing a diagnostic composition for the detection
CC of the disease mentioned above. The antibody is used for the
CC disintegration of beta-amyloid plaques or for passive immunisation
CC against beta-amyloid plaque formation. In particular, the disease is
CC dementia, Alzheimer's disease, motor neuropathy, Down's syndrome,
CC Creutzfeldt Jacob disease, hereditary cerebral haemorrhage with
CC amyloidosis Dutch type, Parkinson's disease, HIV-related dementia,
CC amyotrophic lateral sclerosis or neuronal disorders related to aging. The
CC present sequence is used in the exemplification of the present invention.
XX
XX Sequence 17 AA;
XX
XX Query Match 74.1%; Score 63; DB 6; Length 17;
XX Best Local Similarity 76.5%; Pred. No. 0.00056;
XX Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
XX
XX 1 VISENGRTINYADSVKG 17
XX ||||| : |||||
XX 1 VISETGKNIYADSVKG 17
XX
XX RESULT 5
XX ID ADA90004
XX ID ADA90004 standard; peptide; 17 AA.
XX
XX ADA90004;
XX
XX 20-NOV-2003 (first entry)
XX
XX Anti-Abeta antibody related amino acid sequence SEQ ID NO:119.
XX
XX antibody molecule; antibody; beta-A4 peptide; Abeta4; neuroprotective;
XX neurotropic; antiparkinsonian; gene therapy; amyloidogenesis;
XX amyloid-plaque formation; beta-amyloid plaque; immunisation; dementia;
XX Alzheimer's disease; motor neuropathy; Down's syndrome;
XX Creutzfeldt Jacob disease; hereditary cerebral haemorrhage; amyloidosis;

CC amyloidosis Dutch type, Parkinson's disease, HIV-related dementia,
CC amyotrophic lateral sclerosis or neuronal disorders related to aging. The
CC present sequence is used in the exemplification of the present invention.
XX
XX
SQ Sequence 17 AA; 74.1%; Score 63; DB 6; Length 17; Indels 0; Gaps 0
Query Match Best Local Similarity 76.5%; Pred. No. 0.00056;
Matches 13; Conservative 1; Mismatches 3;
QY 1 VISENGRTINVADSVK 17
| | | | | : | | | | |
DB 1 VISETGKNIYADSVK 17
| | | | | : | | | | |
RESULT 7
ADA90434
ID ADA90434 standard; peptide; 17 AA.
XX
XX AC
ADA90434;
XX
DT 20-NOV-2003 (first entry)
XX
DE MS-Roche #3.4.H7 H-CDR2 amino acid sequence.
XX
XX antibody molecule; antibody; beta-A4 peptide; Abeta4; neuroprotective;
KW neotropic; antiparkinsonian; gene therapy; amyloidogenesis;
KW amyloid-plaque formation; beta-amyloid plaque; immunisation; dementia;
KW Alzheimer's disease; motor neuropathy; Down's syndrome;
KW Creutzfeldt Jacob disease; hereditary cerebral haemorrhage; amyloidosis;
KW Parkinson's disease; HIV-related dementia; amyotrophic lateral sclerosis;
KW neuronal disorder; aging.
XX
XX Synthetic.
OS
OS Homo sapiens.
XX
XX
PN WO2003070760-A2.
XX
XX 28-AUG-2003.
XX
XX 20-FEB-2003; 2003WO-EP001759.
XX
XX 20-FEB-2002; 2002EP-00003844.
XX
XX (HOFF) HOFFMANN LA ROCHE & CO AG F.
XX (MORP-) MORPHOSYS AG.
XX
XX Bardroff M, Bohrmann B, Brockhaus M, Huber W, Kretzschmar T;
PI Loehning C, Loetscher H, Nordstedt C, Rothe C;
XX WPI; 2003-663848/62.
XX
XX New antibody molecule capable of specifically recognizing two regions of
PT the beta-A4 peptide, useful for diagnosing, preventing or treating
PT diseases associated with amyloidogenesis or amyloid-plaque formation
PT (e.g. dementia).
XX
XX Example 5; Page 64; 312pp; English.
XX
XX The present invention describes an antibody molecule (I) capable of
CC specifically recognising two regions of the beta-A4 peptide/Abeta4. The
CC first region comprises the amino acid sequence Ala-Glu-Phe-Arg-His-Asp-
CC Ser-Gly-Tyr ADA89886 or its fragment, and the second region comprises the
CC amino acid sequence Val-His-His-Gln-Lys-Leu-Val-Phe-Phe-Ala-Glu-Asp-Val-
CC Gly ADA9887 or its fragment. Also described: (1) a nucleic acid molecule
CC encoding (I); (2) a vector comprising the nucleic acid of (1); (3) a host
CC cell comprising the vector of (2); (4) preparing (I), comprising
CC culturing the host cell of (3) under conditions that allow synthesis of
CC (I) and recovering (I) from the culture; (5) a composition comprising (I)
CC or an antibody molecule produced by method (4); (6) a kit comprising (I),
CC nucleic acid of (1), vector of (2) or host cell of (3); (7) optimising
CC (I); (8) testing the resulting Fab optimisation library by panning
CC against Abeta/Abeta4; (9) identifying optimised clones; (10) expressing

of selected, optimised clones; (11) preparing a pharmaceutical composition, comprising optimisation of (1), and formulating the optimised antibody/antibody molecule with a carrier; and (12) a pharmaceutical composition prepared by method (8). (1) has neuroprotective, nootropic and antiparkinsonian activities, and can be used in gene therapy. The antibody molecule (1), nucleic acid molecule, vector or host is useful in preparing a pharmaceutical composition for the prevention and/or treatment of a disease associated with amyloidogenesis and/or amyloid-plaque formation. The antibody molecule may also be used in preparing a diagnostic composition for the detection of the disease mentioned above. The antibody is used for the disintegration of beta-amyloid plaques or for passive immunisation against beta-amyloid plaque formation. In particular, the disease is dementia, Alzheimer's disease, motor neuropathy, Down's syndrome, Creutzfeldt Jacob disease, hereditary cerebral haemorrhage with amyloidosis Dutch type, Parkinson's disease, HIV-related dementia, amyotrophic lateral sclerosis or neuronal disorders related to aging. The present sequence is used in the exemplification of the present invention.

XX Sequence 17 AA;

Query Match 74.1%; Score 63; DB 6; Length 17;
Best Local Similarity 76.5%; Pred. No. 0.00056;
Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VISENGRTINVADSVKG 17

DB 1 VISETGKNIYADSVKG 17

RESULT 8

ADA91199
ID ADA91199 standard; peptide; 17 AA.

XX AC ADA91199;

XX DT 20-NOV-2003 (first entry)

XX DE MS-R Fab/antibody related peptide #247.

XX antibody molecule; antibody; beta-A4 peptide; Abeta4; neuroprotective;
KW nootropic; antiparkinsonian; gene therapy; amyloidogenesis;
KW amyloid-plaque formation; beta-amyloid plaque; immunisation; dementia;
KW Alzheimer's disease; motor neuropathy; Down's syndrome;
KW Creutzfeldt Jacob disease; hereditary cerebral haemorrhage; amyloidosis;
KW Parkinson's disease; HIV-related dementia; amyotrophic lateral sclerosis;
KW neuronal disorder; aging.

XX Synthetic.

XX OS Homo sapiens.

XX PN WO2003070760-A2.

XX PD 28-AUG-2003.

XX PF 20-FEB-2003; 2003WO-EP001759.

XX PR 20-FEB-2002; 2002EP-00003844.

XX PA (HOPF) HOFFMANN LA ROCHE & CO AG F.

XX PA (MORP-) MORPHOSYS AG.

XX PI Bardroff M, Bohrman B, Brockhaus M, Huber W, Kretzschmar T;

XX PI Loehning C, Loetscher H, Nordstedt C, Rothe C;

XX DR WPI; 2003-663848/62.

XX New antibody molecule capable of specifically recognizing two regions of
PT the beta-A4 peptide, useful for diagnosing, preventing or treating
PT diseases associated with amyloidogenesis or amyloid-plaque formation
PT (e.g. dementia).

XX Example 13; Page 92; 312pp; English.

XX The present invention describes an antibody molecule (1) capable of
CC specifically recognising two regions of the beta-A4 peptide/Abeta4. The
CC first region comprises the amino acid sequence Ala-Glu-Phe-Arg-His-Asp-
CC Ser-Gly-Tyr ADA9886 or its fragment, and the second region comprises the
CC amino acid sequence Val-His-Gln-Lys-Leu-Val-Phe-Phe-Ala-Glu-Asp-Val-
CC Gly ADA9887 or its fragment. Also described: (1) a nucleic acid molecule
CC encoding (1); (2) a vector comprising the nucleic acid of (1); (3) a host
CC cell comprising the vector of (2); (4) preparing (1), comprising
CC culturing the host cell of (3) under conditions that allow synthesis of
CC (1) and recovering (1) from the culture; (5) a composition comprising (1)
CC or an antibody molecule produced by method (4); (6) a kit comprising (1),
CC nucleic acid of (1), vector of (2) or host cell of (3); (7) optimising
CC (1); (8) testing the resulting Fab optimisation library by panning
CC against Abeta4/Abeta4; (9) identifying optimised clones; (10) expressing
CC of selected, optimised clones; (11) preparing a pharmaceutical
CC composition, comprising optimisation of (1), and formulating the
CC optimised antibody/antibody molecule with a carrier; and (12) a
CC pharmaceutical composition prepared by method (8). (1) has
CC neuroprotective, nootropic and antiparkinsonian activities, and can be
CC used in gene therapy. The antibody molecule (1), nucleic acid molecule,
CC vector or host is useful in preparing a pharmaceutical composition for
CC the prevention and/or treatment of a disease associated with
CC amyloidogenesis and/or amyloid-plaque formation. The antibody molecule
CC may also be used in preparing a diagnostic composition for the detection
CC of the disease mentioned above. The antibody is used for the
CC disintegration of beta-amyloid plaques or for passive immunisation
CC against beta-amyloid plaque formation. In particular, the disease is
CC dementia, Alzheimer's disease, motor neuropathy, Down's syndrome,
CC Creutzfeldt Jacob disease, hereditary cerebral haemorrhage with
CC amyloidosis Dutch type, Parkinson's disease, HIV-related dementia,
CC amyotrophic lateral sclerosis or neuronal disorders related to aging. The
CC present sequence is used in the exemplification of the present invention.

XX Sequence 17 AA;

Query Match 74.1%; Score 63; DB 6; Length 17;
Best Local Similarity 76.5%; Pred. No. 0.00056;

Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VISENGRTINVADSVKG 17

DB 1 VISETGKNIYADSVKG 17

RESULT 9

ADA91301

ID ADA91301 standard; peptide; 17 AA.

XX AC ADA91301;

XX DT 20-NOV-2003 (first entry)

XX DE MS-R Fab/antibody related peptide #349.

XX antibody molecule; antibody; beta-A4 peptide; Abeta4; neuroprotective;
KW nootropic; antiparkinsonian; gene therapy; amyloidogenesis;
KW amyloid-plaque formation; beta-amyloid plaque; immunisation; dementia;
KW Alzheimer's disease; motor neuropathy; Down's syndrome;
KW Creutzfeldt Jacob disease; hereditary cerebral haemorrhage; amyloidosis;
KW Parkinson's disease; HIV-related dementia; amyotrophic lateral sclerosis;
KW neuronal disorder; aging.

XX Synthetic.

XX OS Homo sapiens.

XX PN WO2003070760-A2.

XX PD 28-AUG-2003.

XX PF 20-FEB-2003; 2003WO-EP001759.

XX PR 20-FEB-2002; 2002EP-00003844.

XX (HOFF) HOFFMANN LA ROCHE & CO AG F.
PA (MORP-) MORPHOSYS AG.
XX Bardroff M, Bohrmann B, Brockhaus M, Huber W, Kretzschmar T;
PI Loehning C, Loetscher H, Nordstedt C, Rothe C;
XX WPI; 2003-663848/62.
XX
XX New antibody molecule capable of specifically recognizing two regions of
PT the beta-A4 peptide, useful for diagnosing, preventing or treating
PT diseases associated with amyloidogenesis or amyloid-plaque formation
PT (e.g. dementia).
XX
XX Example 13; Page 93; 312pp; English.
XX
XX The present invention describes an antibody molecule (I) capable of
CC specifically recognising two regions of the beta-A4 peptide/Abeta4. The
CC first region comprises the amino acid sequence Ala-Glu-Phe-Arg-His-Asp-
CC Ser-Gly-Tyr ADA89886 or its fragment, and the second region comprises the
CC amino acid sequence Val-His-His-Gln-Lys-Leu-Phe-Phe-Ala-Glu-Asp-Val-
CC Gly ADA89887 or its fragment. Also described: (1) a nucleic acid molecule
CC encoding (1); (2) a vector comprising the nucleic acid of (1); (3) a host
CC cell comprising the vector of (2); (4) preparing (1), comprising
CC culturing the host cell of (3) under conditions that allow synthesis of
CC (I) and recovering (1) from the culture; (5) a composition comprising (1)
CC or an antibody molecule produced by method (4); (6) a kit comprising (1),
CC nucleic acid of (1), vector of (2) or host cell of (3); (7) optimising
CC (1); (8) testing the resulting Fab optimisation library by panning
CC against Abeta/Abeta4; (9) identifying optimised clones; (10) expressing
CC of selected, optimised clones; (11) preparing a pharmaceutical
CC composition, comprising optimisation of (1), and formulating the
CC optimised antibody/antibody molecule with a carrier; and (12) a
CC pharmaceutical composition prepared by method (8). (1) has
CC neuroprotective, neurotropic and antiparkinsonian activities, and can be
CC used in gene therapy. The antibody molecule (I), nucleic acid molecule,
CC vector or host is useful in preparing a pharmaceutical composition for
CC the prevention and/or treatment of a disease associated with
CC amyloidogenesis and/or amyloid-plaque formation. The antibody molecule
CC may also be used in preparing a diagnostic composition for the detection
CC of the disease mentioned above. The antibody is used for the
CC disintegration of beta-amyloid plaques or for passive immunisation
CC against beta-amyloid plaque formation. In particular, the disease is
CC dementia, Alzheimer's disease, motor neuropathy, Down's syndrome,
CC Creutzfeldt Jacob disease, hereditary cerebral haemorrhage with
CC amyloidosis Dutch type, Parkinson's disease, HIV-related dementia,
CC amyotrophic lateral sclerosis or neuronal disorders related to aging. The
CC present sequence is used in the exemplification of the present invention.
XX Sequence 17 AA;
XX
XX Query Match 74.1%; Score 63; DB 6; Length 17;
XX Best Local Similarity 76.5%; Pred. No. 0.00056;
XX Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
XX
XX QY 1 VISEGRTINYADSVKVG 17
XX |||||:|||||||
XX Db 1 VISEGKNIYADSVKVG 17
XX
XX RESULT 10
XX ADA90190
XX ID ADA90190 standard; protein; 120 AA.
XX
XX AC ADA90190;
XX
XX XX
XX DT 20-NOV-2003 (first entry)
XX
XX DE Anti-Abeta antibody related amino acid sequence SEQ ID NO:305.
XX
XX antibody molecule; antibody; beta-A4 peptide; Abeta4; neuroprotective;
XX neurotropic; antiparkinsonian; gene therapy; amyloidogenesis;
XX amyloid-plaque formation; beta-amyloid plaque; immunisation; dementia;
XX

KW Alzheimer's disease; motor neuropathy; Down's syndrome;
KW Creutzfeldt Jacob disease; hereditary cerebral haemorrhage; amyloidosis;
KW Parkinson's disease; HIV-related dementia; amyotrophic lateral sclerosis;
KW neuronal disorder; aging.
XX Synthetic.
OS Homo sapiens.
XX WO2003070760-A2.
XX
XX 28-AUG-2003.
PD
XX 20-FEB-2003; 2003WO-EP001759.
PF
XX 20-FEB-2002; 2002EP-00003844.
PR
XX (HOFF) HOFFMANN LA ROCHE & CO AG F.
PA (MORP-) MORPHOSYS AG.
XX
XX Bardroff M, Bohrmann B, Brockhaus M, Huber W, Kretzschmar T;
PI Loehning C, Loetscher H, Nordstedt C, Rothe C;
XX WPI; 2003-663848/62.
DR
XX New antibody molecule capable of specifically recognizing two regions of
PT the beta-A4 peptide, useful for diagnosing, preventing or treating
PT diseases associated with amyloidogenesis or amyloid-plaque formation
PT (e.g. dementia).
XX
XX Example 16; Page 272-273; 312pp; English.
XX
XX The present invention describes an antibody molecule (I) capable of
CC specifically recognising two regions of the beta-A4 peptide/Abeta4. The
CC first region comprises the amino acid sequence Ala-Glu-Phe-Arg-His-Asp-
CC Ser-Gly-Tyr ADA89886 or its fragment, and the second region comprises the
CC amino acid sequence Val-His-His-Gln-Lys-Leu-Phe-Phe-Ala-Glu-Asp-Val-
CC Gly ADA89887 or its fragment. Also described: (1) a nucleic acid molecule
CC encoding (1); (2) a vector comprising the nucleic acid of (1); (3) a host
CC cell comprising the vector of (2); (4) preparing (1), comprising
CC culturing the host cell of (3) under conditions that allow synthesis of
CC (I) and recovering (1) from the culture; (5) a composition comprising (1)
CC or an antibody molecule produced by method (4); (6) a kit comprising (1),
CC nucleic acid of (1), vector of (2) or host cell of (3); (7) optimising
CC (1); (8) testing the resulting Fab optimisation library by panning
CC against Abeta/Abeta4; (9) identifying optimised clones; (10) expressing
CC of selected, optimised clones; (11) preparing a pharmaceutical
CC composition, comprising optimisation of (1), and formulating the
CC optimised antibody/antibody molecule with a carrier; and (12) a
CC pharmaceutical composition prepared by method (8). (1) has
CC neuroprotective, neurotropic and antiparkinsonian activities, and can be
CC used in gene therapy. The antibody molecule (I), nucleic acid molecule,
CC vector or host is useful in preparing a pharmaceutical composition for
CC the prevention and/or treatment of a disease associated with
CC amyloidogenesis and/or amyloid-plaque formation. The antibody molecule
CC may also be used in preparing a diagnostic composition for the detection
CC of the disease mentioned above. The antibody is used for the
CC disintegration of beta-amyloid plaques or for passive immunisation
CC against beta-amyloid plaque formation. In particular, the disease is
CC dementia, Alzheimer's disease, motor neuropathy, Down's syndrome,
CC Creutzfeldt Jacob disease, hereditary cerebral haemorrhage with
CC amyloidosis Dutch type, Parkinson's disease, HIV-related dementia,
CC amyotrophic lateral sclerosis or neuronal disorders related to aging. The
CC present sequence is used in the exemplification of the present invention.
XX Sequence 120 AA;
XX
XX Query Match 74.1%; Score 63; DB 6; Length 120;
XX Best Local Similarity 76.5%; Pred. No. 0.0059;
XX Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
XX
XX QY 1 VISEGRTINYADSVKVG 17
XX |||||:|||||||
XX Db 48 VISEGKNIYADSVKVG 64
XX

CC	amyloidosis Dutch type, Parkinson's disease, HIV-related dementia.
CC	anyotrophic lateral sclerosis or neuronal disorders related to aging. The
CC	present sequence is used in the exemplification of the present invention.
XX	
SQ	Sequence 120 AA;
Query Match	74.1%; Score 63; DB 6; Length 120;
Best Local Similarity	76.5%; Pred. No. 0.0059;
Matches 13; Conservative	1; Mismatches 3; Indels 0; Gaps 0;
QY	1 VISENGRTINYADSVK 17
	:
DB	48 VISETGKNVIYADSVK 64
RESULT 12	
ADA90010	ID ADA90010 standard; peptide; 17 AA.
XX	
AC	ADA90010;
XX	
DT	20-NOV-2003 (first entry)
XX	
DE	Anti-Abeta antibody related amino acid sequence SEQ ID NO:125.
XX	
KW	antibody molecule; antibody; beta-A4 peptide; Abeta4; neuroprotective;
KW	nootropic; antiparkinsonian; gene therapy; amyloidogenesis;
KW	amyloid-plaque formation; beta-amyloid plaque; immunisation; dementia;
KW	Alzheimer's disease; motor neuropathy; Down's syndrome;
KW	Creutzfeldt Jacob disease; hereditary cerebral haemorrhage; amyloidosis;
KW	Parkinson's disease; HIV-related dementia; amyotrophic lateral sclerosis;
KW	neuronal disorder; aging.
XX	
OS	Synthetic.
OS	Homo sapiens.
XX	
FN	WO2003070760-A2.
XX	
PD	28-AUG-2003.
XX	
PF	20-FEB-2003; 2003WO-EF001759.
XX	
PR	20-FEB-2002; 2002EP-00003844.
XX	
PA	(HOFF) HOFFMANN LA ROCHE & CO AG F.
PA	(MORP-) MORPHOSYS AG.
XX	
PI	Bardroff M, Bohrmann B, Brockhaus M, Huber W, Kretzschmar T;
PI	Loehning C, Loetscher H, Nordstedt C, Rothe C;
XX	
DR	WPI; 2003-663848/62.
XX	
PT	New antibody molecule capable of specifically recognizing two regions of
PT	the beta-A4 peptide, useful for diagnosing, preventing or treating
PT	diseases associated with amyloidogenesis or amyloid-plaque formation
XX	(e.g. dementia).
XX	
PS	Disclosure; Page 191; 312pp; English.
XX	
CC	The present invention describes an antibody molecule (I) capable of
CC	specifically recognising two regions of the beta-A4 peptide/Abeta4. The
CC	first region comprises the amino acid sequence Ala-Glu-Phe-Arg-His-Asp-
CC	Ser-Gly-Tyr ADAR9886 or its fragment, and the second region comprises the
CC	amino acid sequence Val-His-His-Gln-Lys-Leu-Val-Phe-Ala-Glu-Asp-Val-
CC	Gly ADAR9887 or its fragment. Also described: (1) a nucleic acid molecule
CC	encoding (1); (2) a vector comprising the nucleic acid of (1); (3) a host
CC	cell comprising the vector of (2); (4) preparing (1), comprising
CC	culturing the host cell of (3) under conditions that allow synthesis of
CC	(1) and recovering (1) from the culture; (5) a composition comprising (1)
CC	or an antibody molecule produced by method (4); (6) a kit comprising (1),
CC	nucleic acid of (1), vector of (2) or host cell of (3); (7) optimising
CC	(1); (8) testing the resulting Fab optimisation library by panning
CC	against Abeta/Abeta4; (9) identifying optimised clones; (10) expressing
CC	of selected, optimised clones; (11) preparing a pharmaceutical
CC	composition, comprising optimisation of (1), and formulating the
CC	optimised antibody/molecule with a carrier; and (12) a
CC	pharmaceutical composition prepared by method (8). (1) has
CC	neuroprotective, nootropic and antiparkinsonian activities, and can be
CC	used in gene therapy. The antibody molecule (I), nucleic acid molecule,
CC	vector or host is useful in preparing a pharmaceutical composition for
CC	the prevention and/or treatment of a disease associated with
CC	amyloidogenesis and/or amyloid-plaque formation. The antibody molecule
CC	may also be used in preparing a diagnostic composition for the detection
CC	of the disease mentioned above. The antibody is used for the
CC	disintegration of beta-amyloid plaques or for passive immunisation
CC	against beta-amyloid plaque formation. In particular, the disease is

CC (1); (8) testing the resulting Fab optimisation library by panning
 CC against Abeta4/beta4; (9) identifying optimised clones; (10) expressing
 CC of selected, optimised clones; (11) preparing a pharmaceutical
 CC composition, comprising optimisation of (1), and formulating the
 CC optimised antibody/antibody molecule with a carrier; and (12) a
 CC pharmaceutical composition prepared by method (8). (1) has
 CC neuroprotective, neurotropic and antiparkinsonian activities, and can be
 CC used in gene therapy. The antibody molecule (1), nucleic acid molecule,
 CC vector or host is useful in preparing a pharmaceutical composition for
 CC the prevention and/or treatment of a disease associated with
 CC amyloidogenesis and/or amyloid-plaque formation. The antibody molecule
 CC may also be used in preparing a diagnostic composition for the detection
 CC of the disease mentioned above. The antibody is used for the
 CC disintegration of beta-amyloid plaques or for passive immunisation
 CC against beta-amyloid plaque formation. In particular, the disease is
 CC dementia, Alzheimer's disease, motor neuropathy, Down's syndrome,
 CC Creutzfeldt Jacob disease, hereditary cerebral haemorrhage with
 CC amyloidosis Dutch type, Parkinson's disease, HIV-related dementia,
 CC amyotrophic lateral sclerosis or neuronal disorders related to aging. The
 CC present sequence is used in the exemplification of the present invention.

XX Sequence 17 AA;

Query Match 72.9%; Score 62; DB 6; Length 17;
 Best Local Similarity 81.2%; Pred. No. 0.00084;
 Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVKG 17

DB 2 ISEQGRNIYADSVKG 17

RESULT 13

ADA91195
 ID ADA91195 standard; peptide; 17 AA.

XX AC ADA91195;

XX DT 20-NOV-2003 (first entry)

XX DE MS-R Fab/antibody related peptide #243.

XX antibody molecule; antibody; beta-A4 peptide; Abeta4; neuroprotective;
 KW neurotropic; antiparkinsonian; gene therapy; amyloidogenesis;
 KW amyloid-plaque formation; beta-amyloid plaque; immunisation; dementia;
 KW Alzheimer's disease; motor neuropathy; Down's syndrome;
 KW Creutzfeldt Jacob disease; hereditary cerebral haemorrhage; amyloidosis;
 KW Parkinson's disease; HIV-related dementia; amyotrophic lateral sclerosis;
 KW neuronal disorder; aging.

XX OS Synthetic.

XX OS Homo sapiens.

XX PN WO2003070760-A2.

XX PD 28-AUG-2003.

XX PF 20-FEB-2003; 2003WO-EP001759.

XX PR 20-FEB-2003; 2002EP-00003844.

XX PA (HOPF) HOPFMANN LA ROCHE & CO AG F.

XX PA (MORP-) MORPHOSYS AG.

XX PI Bardroff M, Bohrmann B, Brockhaus M, Huber W, Kretzschmar T;

PI Loehning C, Loetscher H, Nordstedt C, Rothe C;

XX DR WPI; 2003-663848/62.

PT New antibody molecule capable of specifically recognising two regions of
 PT the beta-A4 peptide, useful for diagnosing, preventing or treating
 PT diseases associated with amyloidogenesis or amyloid-plaque formation
 PT (e.g. dementia).

XX Example 13; Page 92; 312pp; English.

CC The present invention describes an antibody molecule (1) capable of
 CC specifically recognising two regions of the beta-A4 peptide/Abeta4. The
 CC first region comprises the amino acid sequence Ala-Glu-Phe-Arg-His-Asp-
 CC Ser-Gly-Tyr ADA89886 or its fragment, and the second region comprises the
 CC amino acid sequence Val-His-His-Gln-Lys-Leu-Val-Phe-Phe-Ala-Glu-Asp-Val-
 CC Gly ADA89887 or its fragment. Also described: (1) a nucleic acid molecule
 CC encoding (1); (2) a vector comprising the nucleic acid of (1); (3) a host
 CC cell comprising the vector of (2); (4) preparing (1), comprising
 CC culturing the host cell of (3) under conditions that allow synthesis of
 CC (1) and recovering (1) from the culture; (5) a composition comprising (1)
 CC or an antibody molecule produced by method (4); (6) a kit comprising (1),
 CC nucleic acid of (1), vector of (2) or host cell of (3); (7) optimising
 CC (1); (8) testing the resulting Fab optimisation library by panning
 CC against Abeta4/beta4; (9) identifying optimised clones; (10) expressing
 CC of selected, optimised clones; (11) preparing a pharmaceutical
 CC composition, comprising optimisation of (1), and formulating the
 CC optimised antibody/antibody molecule with a carrier; and (12) a
 CC pharmaceutical composition prepared by method (8). (1) has
 CC neuroprotective, neurotropic and antiparkinsonian activities, and can be
 CC used in gene therapy. The antibody molecule (1), nucleic acid molecule,
 CC vector or host is useful in preparing a pharmaceutical composition for
 CC the prevention and/or treatment of a disease associated with
 CC amyloidogenesis and/or amyloid-plaque formation. The antibody molecule
 CC may also be used in preparing a diagnostic composition for the detection
 CC of the disease mentioned above. The antibody is used for the
 CC disintegration of beta-amyloid plaques or for passive immunisation
 CC against beta-amyloid plaque formation. In particular, the disease is
 CC dementia, Alzheimer's disease, motor neuropathy, Down's syndrome,
 CC Creutzfeldt Jacob disease, hereditary cerebral haemorrhage with
 CC amyloidosis Dutch type, Parkinson's disease, HIV-related dementia,
 CC amyotrophic lateral sclerosis or neuronal disorders related to aging. The
 CC present sequence is used in the exemplification of the present invention.

XX SQ Sequence 17 AA;

Query Match 72.9%; Score 62; DB 6; Length 17;
 Best Local Similarity 76.5%; Pred. No. 0.00084;

Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VISENGRTINYADSVKG 17

DB 1 VISQGRKIYADSVKG 17

RESULT 14

ADA91198
 ID ADA91198 standard; peptide; 17 AA.

XX AC ADA91198;

XX DT 20-NOV-2003 (first entry)

XX DE MS-R Fab/antibody related peptide #246.

XX antibody molecule; antibody; beta-A4 peptide; Abeta4; neuroprotective;
 KW neurotropic; antiparkinsonian; gene therapy; amyloidogenesis;
 KW amyloid-plaque formation; beta-amyloid plaque; immunisation; dementia;
 KW Alzheimer's disease; motor neuropathy; Down's syndrome;
 KW Creutzfeldt Jacob disease; hereditary cerebral haemorrhage; amyloidosis;
 KW Parkinson's disease; HIV-related dementia; amyotrophic lateral sclerosis;
 KW neuronal disorder; aging.

XX OS Synthetic.

XX OS Homo sapiens.

XX PN WO2003070760-A2.

XX PD 28-AUG-2003.

XX PF 20-FEB-2003; 2003WO-EP001759.

XX PR 20-FEB-2002; 2002EP-00003844.
 XX PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
 XX PA (MORP-) MORPHOSYS AG.
 XX PI Bardroff M, Bohrmann B, Brockhaus M, Huber W, Kretzschmar T;
 XX PI Loehning C, Loetscher H, Nordstedt C, Rothe C;
 XX DR WPI; 2003-663848/62.
 XX XX
 XX PT New antibody molecule capable of specifically recognizing two regions of
 XX PT the beta-A4 peptide, useful for diagnosing, preventing or treating
 XX PT diseases associated with amyloidogenesis or amyloid-plaque formation
 XX PT (e.g. dementia).
 XX XX
 XX PS Example 13; Page 92; 312pp; English.
 XX CC The present invention describes an antibody molecule (I) capable of
 XX CC specifically recognising two regions of the beta-A4 peptide/Abeta4. The
 XX CC first region comprises the amino acid sequence Ala-Glu-Phe-Arg-His-Asp-
 XX CC Ser-Gly-Tyr ADA89886 or its fragment, and the second region comprises the
 XX CC amino acid sequence Val-His-His-Gln-Lys-Leu-Val-Phe-Phe-Ala-Glu-Asp-Val-
 XX CC Gly ADA89887 or its fragment. Also described: (1) a nucleic acid molecule
 XX CC encoding (1); (2) a vector comprising the nucleic acid of (1); (3) a host
 XX CC cell comprising the vector of (2); (4) preparing (1), comprising
 XX CC culturing the host cell of (3) under conditions that allow synthesis of
 XX CC (1) and recovering (1) from the culture; (5) a composition comprising (1)
 XX CC or an antibody molecule produced by method (4); (6) a kit comprising (1),
 XX CC nucleic acid of (1), vector of (2) or host cell of (3); (7) optimising
 XX CC (1); (8) testing the resulting Fab optimisation library by panning
 XX CC against Abeta/Abeta4; (9) identifying optimised clones; (10) expressing
 XX CC of selected, optimised clones; (11) preparing a pharmaceutical
 XX CC composition, comprising optimisation of (1), and formulating the
 XX CC optimised antibody/antibody molecule with a carrier; and (12) a
 XX CC pharmaceutical composition prepared by method (8). (1) has
 XX CC neuroprotective, neurotropic and antiparkinsonian activities, and can be
 XX CC used in gene therapy. The antibody molecule (I), nucleic acid molecule,
 XX CC vector or host is useful in preparing a pharmaceutical composition for
 XX CC the prevention and/or treatment of a disease associated with
 XX CC amyloidogenesis and/or amyloid-plaque formation. The antibody molecule
 XX CC may also be used in preparing a diagnostic composition for the detection
 XX CC of the disease mentioned above. The antibody is used for the
 XX CC disintegration of beta-amyloid plaques or for passive immunisation
 XX CC against beta-amyloid plaque formation. In particular, the disease is
 XX CC dementia, Alzheimer's disease, motor neuropathy, Down's syndrome,
 XX CC Creutzfeldt Jacob disease, hereditary cerebral haemorrhage with
 XX CC amyloidosis Dutch type, Parkinson's disease, HIV-related dementia,
 XX CC amyotrophic lateral sclerosis or neuronal disorders related to aging. The
 XX CC present sequence is used in the exemplification of the present invention.
 XX SQ Sequence 17 AA;
 Query Match 72.9%; Score 62; DB 6; Length 17;
 Best Local Similarity 76.5%; Pred. No. 0.00084;
 Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 1 VISENGRTINYADSVK 17
 Db 1 VISQTKRIYADSVK 17
 RESULT 15
 ADA90000
 ID ADA90000 standard; peptide; 17 AA.
 XX AC ADA90000;
 XX XX
 XX DT 20-NOV-2003 (first entry)
 XX DE Anti-Abeta antibody related amino acid sequence SEQ ID NO:115.
 XX XX
 XX KW antibody molecule; antibody; beta-A4 peptide; Abeta4; neuroprotective;

KW KW nootropic; antiparkinsonian; gene therapy; amyloidogenesis;
 KW KW amyloid-plaque formation; beta-amyloid plaque; immunisation; dementia;
 KW KW Alzheimer's disease; motor neuropathy; Down's syndrome;
 KW KW Creutzfeldt Jacob disease; hereditary cerebral haemorrhage; amyloidosis;
 KW KW Parkinson's disease; HIV-related dementia; amyotrophic lateral sclerosis;
 KW KW neuronal disorder; aging.
 XX XX Synthetic.
 OS OS Homo sapiens.
 XX XX WO2003070760-A2.
 XX PN 28-AUG-2003.
 XX PD 20-FEB-2003; 2003WO-EP001759.
 XX PF 20-FEB-2002; 2002EP-00003844.
 XX PR (HOFF) HOFFMANN LA ROCHE & CO AG F.
 XX PA (MORP-) MORPHOSYS AG.
 XX PI Bardroff M, Bohrmann B, Brockhaus M, Huber W, Kretzschmar T;
 XX PI Loehning C, Loetscher H, Nordstedt C, Rothe C;
 XX XX WPI; 2003-663848/62.
 XX DR
 XX XX New antibody molecule capable of specifically recognizing two regions of
 XX PT the beta-A4 peptide, useful for diagnosing, preventing or treating
 XX PT diseases associated with amyloidogenesis or amyloid-plaque formation
 XX PT (e.g. dementia).
 XX XX
 XX PS Disclosure; Page 188; 312pp; English.
 XX CC The present invention describes an antibody molecule (I) capable of
 XX CC specifically recognising two regions of the beta-A4 peptide/Abeta4. The
 XX CC first region comprises the amino acid sequence Ala-Glu-Phe-Arg-His-Asp-
 XX CC Ser-Gly-Tyr ADA89886 or its fragment, and the second region comprises the
 XX CC amino acid sequence Val-His-His-Gln-Lys-Leu-Val-Phe-Phe-Ala-Glu-Asp-Val-
 XX CC Gly ADA89887 or its fragment. Also described: (1) a nucleic acid molecule
 XX CC encoding (1); (2) a vector comprising the nucleic acid of (1); (3) a host
 XX CC cell comprising the vector of (2); (4) preparing (1), comprising
 XX CC culturing the host cell of (3) under conditions that allow synthesis of
 XX CC (1) and recovering (1) from the culture; (5) a composition comprising (1)
 XX CC or an antibody molecule produced by method (4); (6) a kit comprising (1),
 XX CC nucleic acid of (1), vector of (2) or host cell of (3); (7) optimising
 XX CC (1); (8) testing the resulting Fab optimisation library by panning
 XX CC against Abeta/Abeta4; (9) identifying optimised clones; (10) expressing
 XX CC of selected, optimised clones; (11) preparing a pharmaceutical
 XX CC composition, comprising optimisation of (1), and formulating the
 XX CC optimised antibody/antibody molecule with a carrier; and (12) a
 XX CC pharmaceutical composition prepared by method (8). (1) has
 XX CC neuroprotective, neurotropic and antiparkinsonian activities, and can be
 XX CC used in gene therapy. The antibody molecule (I), nucleic acid molecule,
 XX CC vector or host is useful in preparing a pharmaceutical composition for
 XX CC the prevention and/or treatment of a disease associated with
 XX CC amyloidogenesis and/or amyloid-plaque formation. The antibody molecule
 XX CC may also be used in preparing a diagnostic composition for the detection
 XX CC of the disease mentioned above. The antibody is used for the
 XX CC disintegration of beta-amyloid plaques or for passive immunisation
 XX CC against beta-amyloid plaque formation. In particular, the disease is
 XX CC dementia, Alzheimer's disease, motor neuropathy, Down's syndrome,
 XX CC Creutzfeldt Jacob disease, hereditary cerebral haemorrhage with
 XX CC amyloidosis Dutch type, Parkinson's disease, HIV-related dementia,
 XX CC amyotrophic lateral sclerosis or neuronal disorders related to aging. The
 XX CC present sequence is used in the exemplification of the present invention.
 XX SQ Sequence 17 AA;
 Query Match 72.9%; Score 62; DB 6; Length 17;
 Best Local Similarity 76.5%; Pred. No. 0.00084;
 Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 1 VISENGRTINYADSVK 17

Query Match 72.9%; Score 62; DB 6; Length 17;
 Best Local Similarity 76.5%; Pred. No. 0.00084;
 Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 1 VISENGRTINYADSVK 17

Db 1 VISQTRKIYADSVKG 17
|||: || | |||||
RESULT 16
ADA91193
ID ADA91193 standard; peptide; 17 AA.
XX
XX
XX ADA91193;
XX
XX 20-NOV-2003 (first entry)
XX
XX MS-R Fab/antibody related peptide #241.
XX
XX antibody molecule; antibody; beta-A4 peptide; Abeta4; neuroprotective;
KW nootropic; antiparkinsonian; gene therapy; amyloidogenesis;
KW amyloid-plaque formation; beta-amyloid plaque; immunisation; dementia;
KW Alzheimer's disease; motor neuropathy; Down's syndrome;
KW Creutzfeldt Jacob disease; hereditary cerebral haemorrhage; amyloidosis;
KW Parkinson's disease; HIV-related dementia; amyotrophic lateral sclerosis;
KW neuronal disorder; aging.
XX
XX Synthetic.
OS Homo sapiens.
XX
XX WO2003070760-A2.
XX
XX 28-AUG-2003.
XX
XX 20-FEB-2003; 2003WO-EP001759.
XX
XX 20-FEB-2002; 2002EP-00003844.
XX
XX (HOFF) HOFFMANN LA ROCHE & CO AG F.
PA (MORP-) MORPHOSYS AG.
XX
XX Bardroff M, Bohrmann B, Brockhaus M, Huber W, Kretzschmar T;
PI Loehning C, Loetscher H, Nordstedt C, Rothe C;
XX WPI; 2003-663848/62.
XX
XX New antibody molecule capable of specifically recognizing two regions of
PT the beta-A4 peptide, useful for diagnosing, preventing or treating
PT diseases associated with amyloidogenesis or amyloid-plaque formation
PT (e.g. dementia).
XX
XX Example 13; Page 92; 312pp; English.
XX
XX The present invention describes an antibody molecule (I) capable of
CC specifically recognising two regions of the beta-A4 peptide/Abeta4. The
CC first region comprises the amino acid sequence Ala-Glu-Phe-Arg-His-Asp-
CC Ser-Gly-Tyr ADA89886 or its fragment, and the second region comprises the
CC amino acid sequence Val-His-His-Gln-Lys-Leu-Val-Phe-Phe-Ala-Glu-Asp-Val-
CC Gly ADA89887 or its fragment. Also described: (1) a nucleic acid molecule
CC encoding (1); (2) a vector comprising the nucleic acid of (1); (3) a host
CC cell comprising the vector of (2); (4) preparing (1), comprising
CC culturing the host cell of (3) under conditions that allow synthesis of
CC (1) and recovering (1) from the culture; (5) a composition comprising (1)
CC or an antibody molecule produced by method (4); (6) a kit comprising (1),
CC nucleic acid of (1), vector of (2) or host cell of (3); (7) optimising
CC (1); (8) testing the resulting Fab optimisation library by panning
CC against Abeta/Abeta4; (9) identifying optimised clones; (10) expressing
CC of selected, optimised clones; (11) preparing a pharmaceutical
CC composition, comprising optimisation of (1), and formulating the
CC optimised antibody molecule with a carrier; and (12) a
CC pharmaceutical composition prepared by method (9). (1) has
CC neuroprotective, nootropic and antiparkinsonian activities, and can be
CC used in gene therapy. The antibody molecule (I), nucleic acid molecule,
CC vector or host is useful in preparing a pharmaceutical composition for
CC the prevention and/or treatment of a disease associated with
CC amyloidogenesis and/or amyloid-plaque formation. The antibody molecule
CC may also be used in preparing a diagnostic composition for the detection
CC of the disease mentioned above. The antibody is used for the

CC disintegration of beta-amyloid plaques or for passive immunisation
CC against beta-amyloid plaque formation. In particular, the disease is
CC dementia, Alzheimer's disease, motor neuropathy, Down's syndrome,
CC Creutzfeldt Jacob disease, hereditary cerebral haemorrhage with
CC amyloidosis butch type, Parkinson's disease, HIV-related dementia,
CC amyotrophic lateral sclerosis or neuronal disorders related to aging. The
XX present sequence is used in the exemplification of the present invention.
XX
XX Sequence 17 AA;
SQ
Query Match 72.9%; Score 62; DB 6; Length 17;
Best Local Similarity 76.5%; Pred. No. 0.00084;
Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 VISENGRTINYADSVKG 17
|||: || | |||||
Db 1 VISQTRKIYADSVKG 17
|||: || | |||||
RESULT 17
ADA90430
ID ADA90430 standard; peptide; 17 AA.
XX
XX ADA90430;
AC
XX
XX 20-NOV-2003 (first entry)
DT
XX
XX MS-Roche #3.4.H3 H-CDR2 amino acid sequence.
DE
XX
XX antibody molecule; antibody; beta-A4 peptide; Abeta4; neuroprotective;
KW nootropic; antiparkinsonian; gene therapy; amyloidogenesis;
KW amyloid-plaque formation; beta-amyloid plaque; immunisation; dementia;
KW Alzheimer's disease; motor neuropathy; Down's syndrome;
KW Creutzfeldt Jacob disease; hereditary cerebral haemorrhage; amyloidosis;
KW Parkinson's disease; HIV-related dementia; amyotrophic lateral sclerosis;
KW neuronal disorder; aging.
XX
XX Synthetic.
OS Homo sapiens.
XX
XX WO2003070760-A2.
XX
XX 28-AUG-2003.
XX
XX 20-FEB-2003; 2003WO-EP001759.
XX
XX 20-FEB-2002; 2002EP-00003844.
XX
XX (HOFF) HOFFMANN LA ROCHE & CO AG F.
PA (MORP-) MORPHOSYS AG.
XX
XX Bardroff M, Bohrmann B, Brockhaus M, Huber W, Kretzschmar T;
PI Loehning C, Loetscher H, Nordstedt C, Rothe C;
XX WPI; 2003-663848/62.
XX
XX New antibody molecule capable of specifically recognizing two regions of
PT the beta-A4 peptide, useful for diagnosing, preventing or treating
PT diseases associated with amyloidogenesis or amyloid-plaque formation
PT (e.g. dementia).
XX
XX Example 5; Page 64; 312pp; English.
XX
XX The present invention describes an antibody molecule (I) capable of
CC specifically recognising two regions of the beta-A4 peptide/Abeta4. The
CC first region comprises the amino acid sequence Ala-Glu-Phe-Arg-His-Asp-
CC Ser-Gly-Tyr ADA89886 or its fragment, and the second region comprises the
CC amino acid sequence Val-His-His-Gln-Lys-Leu-Val-Phe-Phe-Ala-Glu-Asp-Val-
CC Gly ADA89887 or its fragment. Also described: (1) a nucleic acid molecule
CC encoding (1); (2) a vector comprising the nucleic acid of (1); (3) a host
CC cell comprising the vector of (2); (4) preparing (1), comprising
CC culturing the host cell of (3) under conditions that allow synthesis of
CC (1) and recovering (1) from the culture; (5) a composition comprising (1)

or an antibody molecule produced by method (4); (6) a kit comprising (1), nucleic acid of (1), vector of (2) or host cell of (3); (7) optimising (1); (8) testing the resulting Fab optimisation library by panning against Abeta4; (9) identifying optimised clones; (10) expressing of selected, optimised clones; (11) preparing a pharmaceutical composition, comprising optimisation of (1), and formulating the optimised antibody/antibody molecule with a carrier; and (12) a pharmaceutical composition prepared by method (8). (1) has neuroprotective, neurotropic and antiparkinsonian activities, and can be used in gene therapy. The antibody molecule (1), nucleic acid molecule, vector or host is useful in preparing a pharmaceutical composition for the prevention and/or treatment of a disease associated with amyloidogenesis and/or amyloid-plaque formation. The antibody molecule may also be used in preparing a diagnostic composition for the detection of the disease mentioned above. The antibody is used for the disintegration of beta-amyloid plaques or for passive immunisation against beta-amyloid plaque formation. In particular, the disease is dementia, Alzheimer's disease, motor neuropathy, Down's syndrome, Creutzfeldt Jacob disease, hereditary cerebral haemorrhage with amyloidosis Dutch type, Parkinson's disease, HIV-related dementia, amyotrophic lateral sclerosis or neuronal disorders related to aging. The present sequence is used in the exemplification of the present invention.

Sequence 17 AA;

Query Match 72.9%; Score 62; DB 6; Length 17;
Best Local Similarity 76.5%; Pred. NO. 0.00084;
Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VISENGRTINYADSVK 17
|||: || |||||
DB 1 VISQGRKIYYADSVK 17

RESULT 18

ADA91196
ID ADA91196 standard; peptide; 17 AA.

AC ADA91196;

XX 20-NOV-2003 (first entry)

DT MS-R Fab/antibody related peptide #244.

XX antibody molecule; antibody; beta-A4 peptide; Abeta4; neuroprotective;
KW neurotropic; antiparkinsonian; gene therapy; amyloidogenesis;
KW amyloid-plaque formation; beta-amyloid plaque; immunisation; dementia;
KW Alzheimer's disease; motor neuropathy; Down's syndrome;
KW Creutzfeldt Jacob disease; hereditary cerebral haemorrhage; amyloidosis;
KW Parkinson's disease; HIV-related dementia; amyotrophic lateral sclerosis;
KW neuronal disorder; aging.

XX Synthetic.

OS Homo sapiens.

XX WO2003070760-A2.

XX 28-AUG-2003.

XX 20-FEB-2003; 2003WO-BP001759.

XX 20-FEB-2002; 2002EP-00003844.

XX (HOFF) HOFFMANN LA ROCHE & CO AG F.
PA (MORP-) MORPHOSYS AG.

XX Bardroff M, Bohrmann B, Brockhaus M, Huber W, Kretzschmar T;
PI Loehning C, Loetscher H, Nordstedt C, Rothe C;

XX WPI; 2003-663848/62.

XX New antibody molecule capable of specifically recognizing two regions of
PT the beta-A4 peptide, useful for diagnosing, preventing or treating

PT diseases associated with amyloidogenesis or amyloid-plaque formation
(e.g. dementia).

XX Example 13; Page 92; 312pp; English.

XX The present invention describes an antibody molecule (1) capable of specifically recognising two regions of the beta-A4 peptide/Abeta4. The first region comprises the amino acid sequence Ala-Glu-Phe-Arg-His-Asp-Ser-Gly-Tyr ADA89886 or its fragment, and the second region comprises the amino acid sequence Val-His-His-Gln-Lys-Leu-Val-Phe-Phe-Ala-Glu-Asp-Val-Gly ADA89887 or its fragment. Also described: (1) a nucleic acid molecule encoding (1); (2) a vector comprising the nucleic acid of (1); (3) a host cell comprising the vector of (2); (4) preparing (1), comprising culturing the host cell of (3) under conditions that allow synthesis of (1) and recovering (1) from the culture; (5) a composition comprising (1) or an antibody molecule produced by method (4); (6) a kit comprising (1), nucleic acid of (1), vector of (2) or host cell of (3); (7) optimising (1); (8) testing the resulting Fab optimisation library by panning against Abeta4; (9) identifying optimised clones; (10) expressing of selected, optimised clones; (11) preparing a pharmaceutical composition, comprising optimisation of (1), and formulating the optimised antibody/antibody molecule with a carrier; and (12) a pharmaceutical composition prepared by method (8). (1) has neuroprotective, neurotropic and antiparkinsonian activities, and can be used in gene therapy. The antibody molecule (1), nucleic acid molecule, vector or host is useful in preparing a pharmaceutical composition for the prevention and/or treatment of a disease associated with amyloidogenesis and/or amyloid-plaque formation. The antibody molecule may also be used in preparing a diagnostic composition for the detection of the disease mentioned above. The antibody is used for the disintegration of beta-amyloid plaques or for passive immunisation against beta-amyloid plaque formation. In particular, the disease is dementia, Alzheimer's disease, motor neuropathy, Down's syndrome, Creutzfeldt Jacob disease, hereditary cerebral haemorrhage with amyloidosis Dutch type, Parkinson's disease, HIV-related dementia, amyotrophic lateral sclerosis or neuronal disorders related to aging. The present sequence is used in the exemplification of the present invention.

XX Sequence 17 AA;

Query Match 72.9%; Score 62; DB 6; Length 17;

Best Local Similarity 76.5%; Pred. NO. 0.00084;

Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VISENGRTINYADSVK 17
|||: || |||||
DB 1 VISQGRKIYYADSVK 17

RESULT 19

ADA90440
ID ADA90440 standard; peptide; 17 AA.

XX ADA90440;

XX 20-NOV-2003 (first entry)

XX MS-Roche #3.4.H13 H-CDR2 amino acid sequence.

XX antibody molecule; antibody; beta-A4 peptide; Abeta4; neuroprotective;
KW neurotropic; antiparkinsonian; gene therapy; amyloidogenesis;
KW amyloid-plaque formation; beta-amyloid plaque; immunisation; dementia;
KW Alzheimer's disease; motor neuropathy; Down's syndrome;
KW Creutzfeldt Jacob disease; hereditary cerebral haemorrhage; amyloidosis;
KW Parkinson's disease; HIV-related dementia; amyotrophic lateral sclerosis;
KW neuronal disorder; aging.

XX Synthetic.

OS Homo sapiens.

XX WO2003070760-A2.

XX 28-AUG-2003.

QY 1 VISENGRTINYADSVKG 17
 DB 48 VISQTRKIYADSVKG 64

RESULT 21
 ADA90184
 ID ADA90184 standard; protein; 120 AA.
 XX
 AC ADA90184;
 DT 20-NOV-2003 (first entry)
 XX
 DE Anti-Abeta antibody related amino acid sequence SEQ ID NO:299.
 XX
 KW antibody molecule; antibody; beta-A4 peptide; Abeta4; neuroprotective;
 KW neurotropic; antiparkinsonian; gene therapy; amyloidogenesis;
 KW amyloid-plaque formation; beta-amyloid plaque; immunisation; dementia;
 KW Alzheimer's disease; motor neuropathy; Down's syndrome;
 KW Creutzfeldt Jacob disease; hereditary cerebral haemorrhage; amyloidosis;
 KW Parkinson's disease; HIV-related dementia; amyotrophic lateral sclerosis;
 KW neuronal disorder; aging.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN WO2003070760-A2.
 XX
 PD 28-AUG-2003.
 XX
 PF 20-FEB-2003; 2003WO-EP001759.
 XX
 PR 20-FEB-2002; 2002EP-00003844.
 XX
 PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
 XX
 PI Bardroff M, Bohrmann B, Brockhaus M, Huber W, Kretzschmar T;
 PI Loehning C, Loetscher H, Nordstedt C, Rothe C;
 XX
 WPI; 2003-663848/62.
 XX
 PT New antibody molecule capable of specifically recognizing two regions of
 PT the beta-A4 peptide, useful for diagnosing, preventing or treating
 PT diseases associated with amyloidogenesis or amyloid-plaque formation
 PT (e.g. dementia).
 XX
 PS Example 16; Page 269; 312pp; English.
 XX
 CC The present invention describes an antibody molecule (I) capable of
 CC specifically recognising two regions of the beta-A4 peptide/Abeta4. The
 CC first region comprises the amino acid sequence Ala-Glu-Phe-Arg-His-Asp-
 CC Ser-Gly-Tyr ADA89886 or its fragment, and the second region comprises the
 CC amino acid sequence Val-His-His-Gln-Lys-Val-Phe-Phe-Ala-Glu-Asp-Val-
 CC Gly ADA89887 or its fragment. Also described: (1) a nucleic acid molecule
 CC encoding (1); (2) a vector comprising the nucleic acid of (1); (3) a host
 CC cell comprising the vector of (2); (4) preparing (1), comprising
 CC culturing the host cell of (3) under conditions that allow synthesis of
 CC (1) and recovering (1) from the culture; (5) a composition comprising (1),
 CC or an antibody molecule produced by method (4); (6) a kit comprising (1),
 CC nucleic acid of (1), vector of (2) or host cell of (3); (7) optimising
 CC (1); (8) testing the resulting Fab optimisation library by panning
 CC against Abeta/Abeta4; (9) identifying optimised clones; (10) expressing
 CC of selected, optimised clones; (11) preparing a pharmaceutical
 CC composition, comprising optimisation of (1), and formulating the
 CC optimised antibody/antibody molecule with a carrier; and (12) a
 CC pharmaceutical composition prepared by method (9). (I) has
 CC neuroprotective, neurotropic and antiparkinsonian activities, and can be
 CC used in gene therapy. The antibody molecule (I), nucleic acid molecule,
 CC vector or host is useful in preparing a pharmaceutical composition for
 CC the prevention and/or treatment of a disease associated with
 CC amyloidogenesis and/or amyloid-plaque formation. The antibody molecule

CC may also be used in preparing a diagnostic composition for the detection
 CC of the disease mentioned above. The antibody is used for the
 CC disintegration of beta-amyloid plaques or for passive immunisation
 CC against beta-amyloid plaque formation. In particular, the disease is
 CC dementia, Alzheimer's disease, motor neuropathy, Down's syndrome,
 CC Creutzfeldt Jacob disease, hereditary cerebral haemorrhage with
 CC amyloidosis Dutch type, Parkinson's disease, HIV-related dementia,
 CC amyotrophic lateral sclerosis or neuronal disorders related to aging. The
 CC present sequence is used in the exemplification of the present invention.
 XX

SQ Sequence 120 AA;

Query Match 72.9%; Score 62; DB 6; Length 120;
 Best Local Similarity 76.5%; Pred. No. 0.0088;
 Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VISENGRTINYADSVKG 17
 DB 48 VISQTRKIYADSVKG 64

RESULT 22

ADA90021
 ID ADA90021 standard; peptide; 17 AA.

XX ADA90021;

XX 20-NOV-2003 (first entry)

DE Anti-Abeta antibody related amino acid sequence SEQ ID NO:136.

KW antibody molecule; antibody; beta-A4 peptide; Abeta4; neuroprotective;
 KW neurotropic; antiparkinsonian; gene therapy; amyloidogenesis;
 KW amyloid-plaque formation; beta-amyloid plaque; immunisation; dementia;
 KW Alzheimer's disease; motor neuropathy; Down's syndrome;
 KW Creutzfeldt Jacob disease; hereditary cerebral haemorrhage; amyloidosis;
 KW Parkinson's disease; HIV-related dementia; amyotrophic lateral sclerosis;
 KW neuronal disorder; aging.

OS Synthetic.

OS Homo sapiens.

PN WO2003070760-A2.

PD 28-AUG-2003.

PF 20-FEB-2003; 2003WO-EP001759.

PR 20-FEB-2002; 2002EP-00003844.

PA (HOFF) HOFFMANN LA ROCHE & CO AG F.

PI (MORP-) MORPHOSYS AG.

PI Bardroff M, Bohrmann B, Brockhaus M, Huber W, Kretzschmar T;

PI Loehning C, Loetscher H, Nordstedt C, Rothe C;

XX WPI; 2003-663848/62.

XX New antibody molecule capable of specifically recognizing two regions of
 PT the beta-A4 peptide, useful for diagnosing, preventing or treating
 PT diseases associated with amyloidogenesis or amyloid-plaque formation
 PT (e.g. dementia).

PS Disclosure; Page 194; 312pp; English.

CC The present invention describes an antibody molecule (I) capable of
 CC specifically recognising two regions of the beta-A4 peptide/Abeta4. The
 CC first region comprises the amino acid sequence Ala-Glu-Phe-Arg-His-Asp-
 CC Ser-Gly-Tyr ADA89886 or its fragment, and the second region comprises the
 CC amino acid sequence Val-His-His-Gln-Lys-Val-Phe-Phe-Ala-Glu-Asp-Val-
 CC Gly ADA89887 or its fragment. Also described: (1) a nucleic acid molecule
 CC encoding (1); (2) a vector comprising the nucleic acid of (1); (3) a host
 CC cell comprising the vector of (2); (4) preparing (1), comprising

CC culturing the host cell of (3) under conditions that allow synthesis of
 CC (I) and recovering (I) from the culture; (5) a composition comprising (I)
 CC or an antibody molecule produced by method (4); (6) a kit comprising (I),
 CC nucleic acid of (1), vector of (2) or host cell of (3); (7) optimising
 CC (I); (8) testing the resulting Fab optimisation library by panning
 CC against Abeta/Abeta4; (9) identifying optimised clones; (10) expressing
 CC of selected, optimised clones; (11) preparing a pharmaceutical
 CC composition, comprising optimisation of (I), and formulating the
 CC optimised antibody/antibody molecule with a carrier; and (12) a
 CC pharmaceutical composition prepared by method (8). (I) has
 CC neuroprotective, neurotropic and antiparkinsonian activities, and can be
 CC used in gene therapy. The antibody molecule (I), nucleic acid molecule,
 CC vector or host is useful in preparing a pharmaceutical composition for
 CC the prevention and/or treatment of a disease associated with
 CC amyloidogenesis and/or amyloid-plaque formation. The antibody molecule
 CC may also be used in preparing a diagnostic composition for the detection
 CC of the disease mentioned above. The antibody is used for the
 CC disintegration of beta-amyloid plaques or for passive immunisation
 CC against beta-amyloid plaque formation. In particular, the disease is
 CC dementia, Alzheimer's disease, motor neuropathy, Down's syndrome,
 CC Creutzfeldt Jacob disease, hereditary cerebral haemorrhage with
 CC amyloidosis Dutch type, Parkinson's disease, HIV-related dementia,
 CC amyotrophic lateral sclerosis or neuronal disorders related to aging. The
 CC present sequence is used in the exemplification of the present invention.

XX Sequence 17 AA;

SQ Query Match 71.8%; Score 61; DB 6; Length 17;
 Best Local Similarity 76.5%; Pred. No. 0.0013;
 Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1-VISENGRTINYADSVKG 17
 ||||| : |||||
 Db 1 VISETGKFIYVADSVKG 17

RESULT 23

ID ADA90580
 AC ADA90580 standard; peptide; 17 AA.

XX ADA90580;

DT 20-NOV-2003 (first entry)

DE MS-Roche #3.6.H3 H-CDR2 amino acid sequence.

XX antibody molecule; antibody; beta-A4 peptide; Abeta4; neuroprotective;
 KW neurotropic; antiparkinsonian; gene therapy; amyloidogenesis;
 KW amyloid-plaque formation; beta-amyloid plaque; immunisation; dementia;
 KW Alzheimer's disease; motor neuropathy; Down's syndrome;
 KW Creutzfeldt Jacob disease; hereditary cerebral haemorrhage;
 KW Parkinson's disease; HIV-related dementia; amyotrophic lateral sclerosis;
 KW neuronal disorder; aging.

XX Synthetic.

OS Homo sapiens.

XX WO2003070760-A2.

XX 28-AUG-2003.

XX 20-FEB-2003; 2003WO-EP001759.

XX 20-FEB-2002; 2002EP-00003844.

XX (HOFF) HOFFMANN LA ROCHE & CO AG F.

PA (MORP-) MORPHOSYS AG.

XX Bardroff M, Bohrmann B, Brockhaus M, Huber W, Kretzschmar T;
 PI Loehning C, Loetscher H, Nordstedt C, Rothe C;
 XX WPI; 2003-663948/62.

PT New antibody molecule capable of specifically recognizing two regions of
 PT the beta-A4 peptide, useful for diagnosing, preventing or treating
 PT diseases associated with amyloidogenesis or amyloid-plaque formation
 PT (e.g. dementia).

XX Example 5; Page 65; 312pp; English.

PS The present invention describes an antibody molecule (I) capable of
 XX specifically recognising two regions of the beta-A4 peptide/Abeta4. The
 CC first region comprises the amino acid sequence Ala-Glu-Phe-Arg-His-Asp-
 CC Ser-Gly-Tyr ADA89886 or its fragment, and the second region comprises the
 CC amino acid sequence Val-His-His-Gln-Lys-Leu-Val-Phe-Phe-Ala-Glu-Asp-Val-
 CC Gly ADA89887 or its fragment. Also described: (1) a nucleic acid molecule
 CC encoding (I); (2) a vector comprising the nucleic acid of (1); (3) a host
 CC cell comprising the vector of (2); (4) preparing (I), comprising
 CC culturing the host cell of (3) under conditions that allow synthesis of
 CC (I) and recovering (I) from the culture; (5) a composition comprising (I)
 CC or an antibody molecule produced by method (4); (6) a kit comprising (I),
 CC nucleic acid of (1), vector of (2) or host cell of (3); (7) optimising
 CC (I); (8) testing the resulting Fab optimisation library by panning
 CC against Abeta/Abeta4; (9) identifying optimised clones; (10) expressing
 CC of selected, optimised clones; (11) preparing a pharmaceutical
 CC composition, comprising optimisation of (I), and formulating the
 CC optimised antibody/antibody molecule with a carrier; and (12) a
 CC pharmaceutical composition prepared by method (8). (I) has
 CC neuroprotective, neurotropic and antiparkinsonian activities, and can be
 CC used in gene therapy. The antibody molecule (I), nucleic acid molecule,
 CC vector or host is useful in preparing a pharmaceutical composition for
 CC the prevention and/or treatment of a disease associated with
 CC amyloidogenesis and/or amyloid-plaque formation. The antibody molecule
 CC may also be used in preparing a diagnostic composition for the detection
 CC of the disease mentioned above. The antibody is used for the
 CC disintegration of beta-amyloid plaques or for passive immunisation
 CC against beta-amyloid plaque formation. In particular, the disease is
 CC dementia, Alzheimer's disease, motor neuropathy, Down's syndrome,
 CC Creutzfeldt Jacob disease, hereditary cerebral haemorrhage with
 CC amyloidosis Dutch type, Parkinson's disease, HIV-related dementia,
 CC amyotrophic lateral sclerosis or neuronal disorders related to aging. The
 CC present sequence is used in the exemplification of the present invention.

XX Sequence 17 AA;

SQ Query Match 71.8%; Score 61; DB 6; Length 17;
 Best Local Similarity 76.5%; Pred. No. 0.0013;
 Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VISENGRTINYADSVKG 17
 ||||| : |||||
 Db 1 VISETGKFIYVADSVKG 17

RESULT 24

AAG62968

ID AAG62968 standard; protein; 121 AA.

XX AAG62968;

XX 01-OCT-2001 (first entry)

DE Amino acid sequence of variable heavy chain fragment of clone DH.

XX Antibody; heavy chain; VH; amyloid protein; blood brain barrier;
 KW endothelial cell; brain cell antigen; inflammation; adhesion molecule;
 KW transferrin receptor; neurological disease; Alzheimer's disease;
 KW prion disease; AIDS-related dementia; epilepsy; brain injury.

OS Homo sapiens.

XX WO200144300-A2.

XX 21-JUN-2001.

XX 27-NOV-2000; 2000WO-GB004501.

XX PR 13-DEC-1999; 99US-0170599P.
 XX PA (CAMB-) CAMBRIDGE ANTIBODY TECHNOLOGY.
 XX PI Webster C, Osbourn J, Ward G, Miller K;
 XX DR WPI; 2001-398131/42.
 XX DR N-PSDB; AAH42412.
 XX PT Mixture or panel of antibodies for selecting specific binding members
 XX PT that cross the blood brain barrier, for use in delivering different
 XX PT molecules and treating neurological diseases.
 XX PS Claim 1; Page 109; 109pp; English.
 XX CC The present sequence represents an antibody variable heavy chain (VH)
 XX CC fragment. The fragment is used to produce a mixture or panel of 5
 XX CC different specific binding members, each comprising an antibody VH and/or
 XX CC VL variable domain and capable, when displayed on the surface of
 XX CC filamentous bacteriophage particles or in the case of a specific binding
 XX CC member comprising the D5 VH and/or VL variable domain when bound to human
 XX CC serum amyloid protein, to pass through a mammalian blood brain barrier
 XX CC (BBB). The panel is useful for the selection of specific binding members
 XX CC with a desired property such as ability to cross BBB, ability to bind
 XX CC endothelial cells or other brain cell antigen, ability to bind areas of
 XX CC inflammation in the brain or BBB breakdown or ability to bind
 XX CC intracellular adhesion molecules and to bind transferrin receptor. The
 XX CC antibodies are useful in diagnosis, prophylaxis and treatment of human or
 XX CC animal body, including neurological diseases, such as Alzheimer's
 XX CC disease, prion disease, AIDS-related dementia, epilepsy and traumatic
 XX CC brain injury and any diseases involving inflammation occurring within the
 XX CC brain or central nervous system
 XX SQ Sequence 121 AA;
 Query Match 71.8%; Score 61; DB 4; Length 121;
 Best Local Similarity 75.0%; Pred. No. 0.013; 2; Mismatches 0; Gaps 0;
 Matches 12; Conservative 2; Indels 0; Gaps 0;
 QY 2 ISENGRTINYADSVKG 17
 ||:|:|:|:|:|
 DB 51 ISNSGNTIHYADSVKG 66
 RESULT 25
 ABP45637
 ID ABP45637 standard; protein; 248 AA.
 XX AC ABP45637;
 XX DT 19-AUG-2002 (first entry)
 XX DE Human BlyS binding scFv SEQ ID 1648.
 XX KW BlyS; B lymphocyte stimulator; TNF superfamily; human; cytostatic;
 XX KW tumour necrosis factor; B cell proliferation; B cell differentiation;
 XX KW immunosuppressive; immunostimulant; immunomodulatory; antirheumatic;
 XX KW antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;
 XX KW systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;
 XX KW common variable immunodeficiency; acquired immunodeficiency syndrome.
 XX OS Homo sapiens.
 XX PN WO200202641-A1.
 XX PD 10-JAN-2002.
 XX PF 15-JUN-2001; 2001WO-US019110.
 XX PR 16-JUN-2000; 2000US-0212210P.
 XX PR 17-OCT-2000; 2000US-0240816P.
 XX PR 16-MAR-2001; 2001US-0276248P.

PR 21-MAR-2001; 2001US-0277379P.
 PR 25-MAY-2001; 2001US-0293499P.
 XX PA (HUMA-) HUMAN GENOME SCI INC.
 XX PA (CAMB-) CAMBRIDGE ANTIBODY TECHNOLOGY.
 XX PI Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;
 XX DR WPI; 2002-114799/15.
 XX PT Antibodies against B Lymphocyte Stimulating polypeptides, useful for the
 XX PT diagnosis and treatment of cancers and immune disorders.
 XX PS Claim 1; Page 2369-2370; 3148pp; English.
 XX CC This invention describes novel antibodies that immunospecifically bind to
 XX CC B Lymphocyte Stimulator (BlyS) polypeptides. BlyS is a member of the
 XX CC tumour necrosis factor (TNF) super family and induces B cell
 XX CC proliferation and differentiation. The antibodies of the invention have
 XX CC cytostatic, immunosuppressive, immunostimulant, immunomodulatory,
 XX CC antirheumatic and antiAIDS activity and can be used in vaccines to
 XX CC inhibit the expression and activity of BlyS. The antibodies bind to BlyS in
 XX CC and so may be used to detect and quantitate the presence of BlyS in
 XX CC biological samples and may be used in this way to diagnose disease
 XX CC associated with aberrant expression of BlyS. They may also be
 XX CC administered to treat diseases associated with aberrant BlyS expression
 XX CC and activity such as cancer, immune, and autoimmune disorders and
 XX CC diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis,
 XX CC immunodeficiency (e.g. common variable immunodeficiency (CVID) and
 XX CC acquired immunodeficiency syndrome (AIDS)). ABP43990-ABP47228 represent
 XX CC the antibodies and fragments of the antibodies described in the method of
 XX CC the invention
 XX SQ Sequence 248 AA;
 Query Match 71.8%; Score 61; DB 5; Length 248;
 Best Local Similarity 64.7%; Pred. No. 0.031;
 Matches 11; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 QY 1 VISENGRTINYADSVKG 17
 ||:|:|:|:|:|
 DB 50 VISDDGNVNVEDSVKG 66

Search completed: September 24, 2004, 01:53:20
 Job time : 113.377 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 01:53:42 ; Search time 28.8679 Seconds
(without alignments)
30.402 Million cell updates/sec

Title: US-10-088-639A-2_COPY_177_193
Perfect score: 85
Sequence: 1 VISENGRTINADVSKG 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database : Issued Patents AA:..
1: /cgn2_6/prodata/2/1aa/5A-COMB.pep:.*
2: /cgn2_6/prodata/2/1aa/5B-COMB.pep:.*
3: /cgn2_6/prodata/2/1aa/6A-COMB.pep:.*
4: /cgn2_6/prodata/2/1aa/6B-COMB.pep:.*
5: /cgn2_6/prodata/2/1aa/PCUS-COMB.pep:.*
6: /cgn2_6/prodata/2/1aa/backfiles1.pep:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	60	70.6	120	2	US-07-934-373C-4
2	60	70.6	120	3	US-08-437-642B-4
3	60	70.6	120	4	US-08-146-206C-4
4	60	70.6	120	4	US-09-705-686-4
5	60	70.6	120	5	PCT-US93-07832-4
6	59	69.4	129	3	US-09-240-274-143
7	57	67.1	22	4	US-08-918-148-16
8	57	67.1	98	2	US-08-665-202-31
9	57	67.1	98	4	US-09-315-574-31
10	57	67.1	117	3	US-08-545-809A-99
11	57	67.1	123	2	US-08-665-202-30
12	57	67.1	123	4	US-09-315-574-30
13	57	67.1	245	4	US-08-918-148-75
14	56	65.9	124	3	US-09-240-274-4
15	54	63.5	124	3	US-09-240-274-2
16	54	63.5	124	3	US-09-240-274-3
17	54	63.5	124	3	US-09-240-274-7
18	53	62.4	116	3	US-08-545-809A-143
19	52	61.2	17	1	US-08-264-093-22
20	52	61.2	118	2	US-08-652-816A-11
21	52	61.2	120	1	US-08-264-093-14
22	52	61.2	179	3	US-08-862-124-2
23	52	61.2	287	3	US-08-862-124-17
24	52	61.2	304	3	US-08-862-124-14
25	51	60.0	98	1	US-08-211-202-116
26	51	60.0	110	1	US-08-211-202-117
27	51	60.0	111	1	US-08-211-202-134

28	51	60.0	114	2	US-08-652-558-52	Sequence 52, Appl
29	51	60.0	117	3	US-08-545-809A-130	Sequence 130, Appl
30	51	60.0	121	4	US-09-553-949-7	Sequence 7, Appl
31	51	60.0	121	4	US-09-530-139-60	Sequence 60, Appl
32	51	60.0	126	1	US-08-478-039-95	Sequence 95, Appl
33	51	60.0	126	1	US-08-476-349A-95	Sequence 95, Appl
34	51	60.0	126	3	US-08-983-607-48	Sequence 48, Appl
35	51	60.0	137	1	US-08-331-398A-61	Sequence 61, Appl
36	51	60.0	137	2	US-08-331-397B-61	Sequence 61, Appl
37	51	60.0	137	2	US-08-759-804A-60	Sequence 60, Appl
38	50	58.8	112	1	US-08-211-202-133	Sequence 133, Appl
39	50	58.8	117	1	US-07-942-245-34	Sequence 34, Appl
40	50	58.8	117	3	US-08-545-809A-119	Sequence 119, Appl
41	50	58.8	118	3	US-08-545-809A-97	Sequence 97, Appl
42	50	58.8	120	2	US-08-958-201-8	Sequence 10, Appl
43	50	58.8	120	2	US-08-958-201-10	Sequence 26, Appl
44	50	58.8	126	3	US-08-983-607-26	Sequence 28, Appl
45	49	57.6	22	4	US-08-918-148-28	Sequence 6, Appl
46	49	57.6	113	3	US-08-974-899-6	Sequence 15, Appl
47	49	57.6	119	4	US-09-648-067A-15	Sequence 21, Appl
48	49	57.6	122	2	US-07-934-373C-21	Sequence 21, Appl
49	49	57.6	122	3	US-08-437-642B-21	Sequence 21, Appl
50	49	57.6	122	4	US-08-146-206C-21	Sequence 21, Appl
51	49	57.6	122	4	US-09-705-686-21	Sequence 21, Appl
52	49	57.6	122	5	PCT-US93-07832-21	Sequence 21, Appl
53	49	57.6	153	4	US-09-472-087-8	Sequence 8, Appl
54	49	57.6	153	4	US-09-472-087-77	Sequence 77, Appl
55	49	57.6	163	4	US-09-472-087-3	Sequence 3, Appl
56	49	57.6	163	4	US-09-472-087-78	Sequence 78, Appl
57	49	57.6	167	4	US-09-472-087-6	Sequence 6, Appl
58	49	57.6	167	4	US-09-472-087-76	Sequence 76, Appl
59	49	57.6	245	4	US-08-918-148-76	Sequence 76, Appl
60	48	56.5	67	1	US-08-162-102C-36	Sequence 36, Appl
61	48	56.5	102	3	US-08-477-347-9	Sequence 9, Appl
62	48	56.5	102	4	US-09-800-908-9	Sequence 9, Appl
63	48	56.5	116	2	US-08-652-816A-14	Sequence 14, Appl
64	48	56.5	117	1	US-07-942-245-36	Sequence 36, Appl
65	48	56.5	117	3	US-08-545-809A-106	Sequence 106, Appl
66	48	56.5	117	4	US-09-025-769B-24	Sequence 24, Appl
67	48	56.5	124	1	US-08-276-852-146	Sequence 146, Appl
68	48	56.5	124	1	US-08-899-575-146	Sequence 146, Appl
69	48	56.5	124	1	US-08-899-575-146	Sequence 146, Appl
70	48	56.5	124	5	PCT-US95-08743-146	Sequence 146, Appl
71	48	56.5	129	1	US-08-162-102C-21	Sequence 21, Appl
72	48	56.5	129	3	US-09-000-088-4	Sequence 4, Appl
73	48	56.5	129	4	US-08-920-100B-21	Sequence 21, Appl
74	48	56.5	129	5	PCT-US93-08786-21	Sequence 21, Appl
75	48	56.5	233	3	US-08-480-753-2	Sequence 2, Appl
76	48	56.5	233	3	US-09-041-889-9	Sequence 9, Appl
77	48	56.5	233	3	US-08-837-058-9	Sequence 9, Appl
78	48	56.5	233	4	US-09-417-264-9	Sequence 9, Appl
79	48	56.5	244	2	US-08-480-753-4	Sequence 4, Appl
80	48	56.5	309	4	US-09-079-029-9	Sequence 9, Appl
81	48	56.5	334	4	US-09-646-028-53	Sequence 53, Appl
82	48	56.5	339	4	US-09-646-028-55	Sequence 55, Appl
83	48	56.5	348	4	US-09-646-028-51	Sequence 51, Appl
84	47	55.3	17	4	US-09-383-667-16	Sequence 16, Appl
85	47	55.3	98	1	US-08-211-202-118	Sequence 118, Appl
86	47	55.3	113	2	US-08-956-047-30	Sequence 30, Appl
87	47	55.3	115	2	US-08-428-197-42	Sequence 42, Appl
88	47	55.3	115	5	PCT-US93-10555-42	Sequence 42, Appl
89	47	55.3	116	1	US-08-211-202-141	Sequence 141, Appl
90	47	55.3	117	3	US-08-545-809A-115	Sequence 115, Appl
91	47	55.3	118	3	US-08-545-809A-145	Sequence 145, Appl
92	47	55.3	119	1	US-08-331-397B-46	Sequence 46, Appl
93	47	55.3	119	2	US-08-331-397B-46	Sequence 46, Appl
94	47	55.3	119	3	US-08-759-804A-46	Sequence 46, Appl
95	47	55.3	119	3	US-09-227-693-46	Sequence 46, Appl
96	47	55.3	120	1	US-08-211-202-135	Sequence 135, Appl
97	47	55.3	120	1	US-07-942-245-35	Sequence 35, Appl
98	47	55.3	123	3	US-08-983-607-38	Sequence 38, Appl
99	47	55.3	128	1	US-08-478-039-96	Sequence 96, Appl
100	47	55.3	128	1	US-08-476-349A-96	Sequence 96, Appl

ATTORNEY/AGENT INFORMATION:
 NAME: Lee, Wendy M.
 REGISTRATION NUMBER: 40,378
 REFERENCE/DOCKET NUMBER: P0709P2C1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 650/225-1994
 TELEFAX: 650/952-9881
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 120 amino acids
 TYPE: Amino Acid
 TOPOLOGY: Linear
 US-08-437-642B-4

Query Match 70.6%; Score 60; DB 3; Length 120;
 Best Local Similarity 76.5%; Pred. No. 0.0087;
 Matches 13; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 VISENGRTINYADSVKG 17
 DB 50 VISENGSDTYADSVKG 66

RESULT 3
 US-08-146-206C-4
 Sequence 4, Application US/08146206C
 Patent No. 6407213
 GENERAL INFORMATION:
 APPLICANT: Carter, Paul J.
 APPLICANT: Presta, Leonard G.
 TITLE OF INVENTION: Method for Making Humanized Antibodies
 NUMBER OF SEQUENCES: 26
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Genentech, Inc.
 STREET: 1 DNA Way
 CITY: South San Francisco
 STATE: California
 COUNTRY: USA
 ZIP: 94080

COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: WinPatIn (Genentech)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/146,206C
 FILING DATE: 17-No. 6407213-1993
 CLASSIFICATION: 530

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/715272
 FILING DATE: 14-JUN-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Lee, Wendy M.
 REGISTRATION NUMBER: 40,378
 REFERENCE/DOCKET NUMBER: P0709P1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 650/225-1994
 TELEFAX: 650/952-9881
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 120 amino acids
 TYPE: Amino Acid
 TOPOLOGY: Linear
 US-08-146-206C-4

Query Match 70.6%; Score 60; DB 4; Length 120;
 Best Local Similarity 76.5%; Pred. No. 0.0087;
 Matches 13; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 VISENGRTINYADSVKG 17
 DB 50 VISENGSDTYADSVKG 66

RESULT 4
 US-09-705-686-4
 Sequence 4, Application US/09705686
 Patent No. 6639055
 GENERAL INFORMATION:
 APPLICANT: Carter, Paul J.
 APPLICANT: Presta, Leonard G.
 TITLE OF INVENTION: Method for Making Humanized Antibodies
 NUMBER OF SEQUENCES: 26
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Genentech, Inc.
 STREET: 1 DNA Way
 CITY: South San Francisco
 STATE: California
 COUNTRY: USA
 ZIP: 94080

COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: WinPatIn (Genentech)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/705,686
 FILING DATE: 02-No. 6639055-2000
 CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/146206
 FILING DATE: 17-NOV-1993
 APPLICATION NUMBER: 07/715272
 FILING DATE: 14-JUN-1991

ATTORNEY/AGENT INFORMATION:
 NAME: Lee, Wendy M.
 REGISTRATION NUMBER: 40,378
 REFERENCE/DOCKET NUMBER: P0709P1D3
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 650/225-1994
 TELEFAX: 650/952-9881

INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 120 amino acids
 TYPE: Amino Acid
 TOPOLOGY: Linear
 SEQUENCE DESCRIPTION: SEQ ID NO: 4:

US-09-705-686-4
 Query Match 70.6%; Score 60; DB 4; Length 120;
 Best Local Similarity 76.5%; Pred. No. 0.0087;
 Matches 13; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 VISENGRTINYADSVKG 17
 DB 50 VISENGSDTYADSVKG 66

RESULT 5
 PCT-US93-07832-4
 Sequence 4, Application PC/TUS9307832
 GENERAL INFORMATION:
 APPLICANT: Genentech, Inc.
 TITLE OF INVENTION: Immunoglobulin Variants
 NUMBER OF SEQUENCES: 40
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Genentech, Inc.
 STREET: 460 Point San Bruno Blvd
 CITY: South San Francisco
 STATE: California
 COUNTRY: USA
 ZIP: 94080
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS

```

; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/07832
; FILING DATE: 19930820
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/715272
; FILING DATE: 14-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/05126
; FILING DATE: 15-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/934373
; FILING DATE: 21-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME:
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER: 709P2PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 120 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
PCT-US93-07832-4

Query Match 70.6%; Score 60; DB 5; Length 120;
Best Local Similarity 76.5%; Pred. No. 0.0087; 4; Indels 0; Gaps 0;
Matches 13; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 VISENGRTINYADSVKG 17
Db 50 VISENGGYTRYADSVKG 66

RESULT 6
US-09-240-274-143
; Sequence 143, Application US/09240274
; Patent No. 6255455
; GENERAL INFORMATION:
; APPLICANT: Siegel, Donald L.
; TITLE OF INVENTION: Rh(D)-BINDING PROTEINS AND MAGNETICALLY ACTIVATED CELL
; TITLE OF INVENTION: SORTING METHOD FOR PRODUCTION THEREOF
; FILE REFERENCE: 09596-42U2
; CURRENT APPLICATION NUMBER: US/09/240,274
; CURRENT FILING DATE: 1999-01-29
; EARLIER APPLICATION NUMBER: 60/081,380
; EARLIER FILING DATE: 1998-04-10
; EARLIER APPLICATION NUMBER: 60/028,550
; EARLIER FILING DATE: 1996-10-11
; NUMBER OF SEQ ID NOS: 224
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 143
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: anti-Rh(D) antibody clone SH20
US-09-240-274-143

Query Match 69.4%; Score 59; DB 3; Length 129;
Best Local Similarity 76.5%; Pred. No. 0.014; 3; Indels 0; Gaps 0;
Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VISENGRTINYADSVKG 17
Db 51 VISYDGTIYADSVKG 67

RESULT 7
US-08-918-148-16
; Sequence 16, Application US/08918148A
; Patent No. 6342220
; GENERAL INFORMATION:
; APPLICANT: Adams, Camellia
; APPLICANT: W.
; APPLICANT: Carter, Paul J.
; APPLICANT: Cardly, Brian M.
; APPLICANT: Gurney, Austin L.
; TITLE OF INVENTION: Agonist Antibodies
; FILE REFERENCE: P0979
; CURRENT APPLICATION NUMBER: US/08/918,148A
; CURRENT FILING DATE: 1997-08-25
; NUMBER OF SEQ ID NOS: 79
; SEQ ID NO 16
; LENGTH: 22
; TYPE: PRT
; ORGANISM: artificial
; FEATURE:
; NAME/KEY: SE5scFv VH CDR2
; LOCATION: 1-22
; OTHER INFORMATION:
US-08-918-148-16

Query Match 67.1%; Score 57; DB 4; Length 22;
Best Local Similarity 75.0%; Pred. No. 0.0041; 3; Indels 0; Gaps 0;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVKG 17
Db 2 ISSSGSTIYADSVKG 17

RESULT 8
US-08-665-202-31
; Sequence 31, Application US/08665202
; Patent No. 5977322
; GENERAL INFORMATION:
; APPLICANT: Marks, James D.
; APPLICANT: Schier, Robert
; TITLE OF INVENTION: No. 5977322el High Affinity Human Antibodies to
; TITLE OF INVENTION: Tumor Antigens
; NUMBER OF SEQUENCES: 141
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/665,202
; FILING DATE: 13-JUN-1996
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/000,238
; FILING DATE: 14-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/000,250
; FILING DATE: 15-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Hunter, Tom
; REGISTRATION NUMBER: 38,498
; REFERENCE/DOCKET NUMBER: 02307E-061410
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 31:
```

SEQUENCE CHARACTERISTICS:
LENGTH: 98 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-665-202-31

Query Match 67.1%; Score 57; DB 2; Length 98;
Best Local Similarity 75.0%; Pred. No. 0.022;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVKG 17
DB 51 ISSSGSTIYADSVKG 66

RESULT 9

US-09-315-574-31
Sequence 31, Application US/09315574
Patent No. 6512097
GENERAL INFORMATION:
APPLICANT: Marks, James D.
APPLICANT: Schier, Robert
TITLE OF INVENTION: No. 6512097el High Affinity Human Antibodies to
TITLE OF INVENTION: Tumor Antigens
NUMBER OF SEQUENCES: 141
CORRESPONDENCE ADDRESS:
ADDRESSEE: Majestic, Parsons, Siebert & Heue P.C.
STREET: Four Embarcadero Center, Suite 1100
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4106

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/315,574
FILING DATE: 20-MAY-99
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/000,238
FILING DATE: 14-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/000,250
FILING DATE: 15-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/665,202
FILING DATE: 13-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Hunter, Tom
REGISTRATION NUMBER: 38,498
REFERENCE/DOCKET NUMBER: 02307E-061411
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 98 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-315-574-31

Query Match 67.1%; Score 57; DB 4; Length 98;
Best Local Similarity 75.0%; Pred. No. 0.022;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVKG 17

DB 51 ISSSGSTIYADSVKG 66

RESULT 10

US-08-545-809A-99
Sequence 99, Application US/08545809A
Patent No. 6096878
GENERAL INFORMATION:
APPLICANT: Honjo, Tasuku
APPLICANT: Matsuda, Fumihiko
TITLE OF INVENTION: HUMAN IMMUNOGLOBULIN VH GENE
TITLE OF INVENTION: SEGMENTS AND DNA FRAGMENTS CONTAINING THE SAME
NUMBER OF SEQUENCES: 145
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson, P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02110-2804

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/545,809A
FILING DATE: 27-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP93/00603
FILING DATE: 10-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Freeman, John W.
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 06501/004001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-542-5070
TELEFAX: 617-542-8906
TELEX: 200154

INFORMATION FOR SEQ ID NO: 99:
SEQUENCE CHARACTERISTICS:
LENGTH: 117 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-545-809A-99

Query Match 67.1%; Score 57; DB 3; Length 117;
Best Local Similarity 75.0%; Pred. No. 0.027;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVKG 17
DB 70 ISSSGSTIYADSVKG 85

RESULT 11

US-08-665-202-30
Sequence 30, Application US/08665202
Patent No. 5977322
GENERAL INFORMATION:
APPLICANT: Marks, James D.
APPLICANT: Schier, Robert
TITLE OF INVENTION: No. 5977322el High Affinity Human Antibodies to
TITLE OF INVENTION: Tumor Antigens
NUMBER OF SEQUENCES: 141
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA

```
;
;
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/665,202
; FILING DATE: 13-JUN-1996
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/000,238
; FILING DATE: 14-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/000,250
; FILING DATE: 15-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Hunter, Tom
; REGISTRATION NUMBER: 38,498
; REFERENCE/DOCKET NUMBER: 02307E-061410
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 123 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-665-202-30

Query Match 67.1%; Score 57; DB 2; Length 123;
Best Local Similarity 75.0%; Pred. No. 0.029;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVKVG 17
Db 51 ISSSGSTIYADSVKVG 66

RESULT 12
US-09-315-574-30
; Sequence 30, Application US/09315574
; Patent No. 6512097
; GENERAL INFORMATION:
; APPLICANT: Marks, James D.
; APPLICANT: Schier, Robert
; TITLE OF INVENTION: No. 6512097el High Affinity Human Antibodies to
; NUMBER OF SEQUENCES: 141
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Majestic, Parsons, Siebert & Haue P.C.
; STREET: Four Embarcadero Center, Suite 1100
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4106
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/315,574
; FILING DATE: 20-MAY-99
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/000,238
; FILING DATE: 14-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/000,250
; FILING DATE: 15-JUN-1995

;
;
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/665,202
; FILING DATE: 13-JUN-1996
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/000,238
; FILING DATE: 14-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/000,250
; FILING DATE: 15-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Hunter, Tom
; REGISTRATION NUMBER: 38,498
; REFERENCE/DOCKET NUMBER: 02307E-061411
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 123 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-315-574-30

Query Match 67.1%; Score 57; DB 4; Length 123;
Best Local Similarity 75.0%; Pred. No. 0.029;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVKVG 17
Db 51 ISSSGSTIYADSVKVG 66

RESULT 13
US-08-918-148-75
; Sequence 75, Application US/08918148A
; Patent No. 6342220
; GENERAL INFORMATION:
; APPLICANT: W.
; APPLICANT: Carter, Paul J.
; APPLICANT: Pendlly, Brian M.
; APPLICANT: Gurney, Austin L.
; TITLE OF INVENTION: Agonist Antibodies
; FILE REFERENCE: P0979
; CURRENT APPLICATION NUMBER: US/08/918,148A
; CURRENT FILING DATE: 1997-08-25
; NUMBER OF SEQ ID NOS: 79
; SEQ ID NO 75
; LENGTH: 245
; TYPE: PRT
; ORGANISM: artificial
; US-08-918-148-75

Query Match 67.1%; Score 57; DB 4; Length 245;
Best Local Similarity 75.0%; Pred. No. 0.062;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVKVG 17
Db 53 ISSSGSTIYADSVKVG 68

RESULT 14
US-09-240-274-4
; Sequence 4, Application US/09240274
; Patent No. 6255455
; GENERAL INFORMATION:
; APPLICANT: Siegel, Donald L.
; TITLE OF INVENTION: Rh(D)-BINDING PROTEINS AND MAGNETICALLY ACTIVATED CELL
; TITLE OF INVENTION: SORTING METHOD FOR PRODUCTION THEREOF
; FILE REFERENCE: 09596-42U2
; CURRENT APPLICATION NUMBER: US/09/240,274
; CURRENT FILING DATE: 1999-01-29
; EARLIER APPLICATION NUMBER: 60/081,380
; EARLIER FILING DATE: 1998-04-10
; EARLIER APPLICATION NUMBER: 60/028,550
; EARLIER FILING DATE: 1996-10-11
```

```
; NUMBER OF SEQ ID NOS: 224
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; TYPE: PRT
; LENGTH: 124
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: anti-Rh(D) chain C04
US-09-240-274-4

Query Match          65.5%; Score 56; DB 3; Length 124;
Best Local Similarity 70.6%; Pred. No. 0.043;
Matches 12; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 VISENGRTINYADSVKG 17
   |||:| |||||
Db 50 VISYDGHKNYADSVKG 66

RESULT 15
US-09-240-274-2
; Sequence 2, Application US/09240274
; Patent No. 6255455
; GENERAL INFORMATION:
; APPLICANT: Siegel, Donald L.
; TITLE OF INVENTION: Rh(D)-BINDING PROTEINS AND MAGNETICALLY ACTIVATED CELL
; FILE REFERENCE: 09596-42U2
; CURRENT APPLICATION NUMBER: US/09/240,274
; CURRENT FILING DATE: 1999-01-29
; EARLIER APPLICATION NUMBER: 60/081,380
; EARLIER FILING DATE: 1998-04-10
; EARLIER APPLICATION NUMBER: 60/028,550
; EARLIER FILING DATE: 1996-10-11
; NUMBER OF SEQ ID NOS: 224
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: anti-Rh(D) chain C01
US-09-240-274-2

Query Match          63.5%; Score 54; DB 3; Length 124;
Best Local Similarity 70.6%; Pred. No. 0.092;
Matches 12; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 VISENGRTINYADSVKG 17
   |||:| |||||
Db 50 VISYDGHKNYADSVKG 66

RESULT 16
US-09-240-274-3
; Sequence 3, Application US/09240274
; Patent No. 6255455
; GENERAL INFORMATION:
; APPLICANT: Siegel, Donald L.
; TITLE OF INVENTION: Rh(D)-BINDING PROTEINS AND MAGNETICALLY ACTIVATED CELL
; FILE REFERENCE: 09596-42U2
; CURRENT APPLICATION NUMBER: US/09/240,274
; CURRENT FILING DATE: 1999-01-29
; EARLIER APPLICATION NUMBER: 60/081,380
; EARLIER FILING DATE: 1998-04-10
; EARLIER APPLICATION NUMBER: 60/028,550
; EARLIER FILING DATE: 1996-10-11
; NUMBER OF SEQ ID NOS: 224
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: anti-Rh(D) chain C03
US-09-240-274-3

Query Match          63.5%; Score 54; DB 3; Length 124;
Best Local Similarity 70.6%; Pred. No. 0.092;
Matches 12; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 VISENGRTINYADSVKG 17
   |||:| |||||
Db 50 VISYDGHKNYADSVKG 66

RESULT 17
US-09-240-274-7
; Sequence 7, Application US/09240274
; Patent No. 6255455
; GENERAL INFORMATION:
; APPLICANT: Siegel, Donald L.
; TITLE OF INVENTION: Rh(D)-BINDING PROTEINS AND MAGNETICALLY ACTIVATED CELL
; FILE REFERENCE: 09596-42U2
; CURRENT APPLICATION NUMBER: US/09/240,274
; CURRENT FILING DATE: 1999-01-29
; EARLIER APPLICATION NUMBER: 60/081,380
; EARLIER FILING DATE: 1998-04-10
; EARLIER APPLICATION NUMBER: 60/028,550
; EARLIER FILING DATE: 1996-10-11
; NUMBER OF SEQ ID NOS: 224
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: anti-Rh(D) chain C10
US-09-240-274-7

Query Match          63.5%; Score 54; DB 3; Length 124;
Best Local Similarity 70.6%; Pred. No. 0.092;
Matches 12; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 VISENGRTINYADSVKG 17
   |||:| |||||
Db 50 VISYDGHKNYADSVKG 66

RESULT 18
US-08-545-809A-143
; Sequence 143, Application US/08545809A
; Patent No. 6096878
; GENERAL INFORMATION:
; APPLICANT: Honjo, Tasuku
; APPLICANT: Matsuda, Fumihiko
; TITLE OF INVENTION: HUMAN IMMUNOGLOBULIN VH GENE
; TITLE OF INVENTION: SEGMENTS AND DNA FRAGMENTS CONTAINING THE SAME
; NUMBER OF SEQUENCES: 145
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/545,809A
; FILING DATE: 27-MAR-1996
```

Gaps 0:

QY 2 ISENGRTINYADSVKG 17
||| : |||||
DB 51 ISSNGGSTYADSVKG 66

RESULT 21
US-08-264-093-14
; Sequence 14, Application US/08264093
; Patent No. 5639863
; GENERAL INFORMATION:
; APPLICANT: Michael D. Dan
; TITLE OF INVENTION: HUMAN MONOCLONAL ANTIBODIES SPECIFIC TO
; TITLE OF INVENTION: CELL CYCLE-INDEPENDENT GLIOMA SURFACE
; TITLE OF INVENTION: ANTIGEN
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ridout & Maybee
; STREET: 2300 Richmond-Adelaide Centre
; STREET: 101 Richmond Street West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 2J7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette - 3.5 inch, 1.4 Mb storage
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: MS-DOS 6.00
; SOFTWARE: ASCII Editor
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/264,093
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA: No. 5639863 applicable
; ATTORNEY/AGENT INFORMATION:
; NAME: Lake, James R.
; REGISTRATION NUMBER: 31081
; REFERENCE/DOCKET NUMBER: NOVOP/106A/7551
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 868-1482
; TELEFAX: (416) 362-0823
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 120 amino acids
; TYPE: amino acid
; STRANDEDNESS: not applicable
; TOPOLOGY: linear
US-08-264-093-14

Query Match 61.2%; Score 52; DB 1; Length 120;
Best Local Similarity 68.8%; Pred. No. 0.19;
Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVKG 17
||| : |||||
DB 44 ISSNGGSTYADSVKG 59

RESULT 22
US-08-862-124-2
; Sequence 2, Application US/08862124
; Patent No. 6207153
; GENERAL INFORMATION:
; APPLICANT: Dan, Michael D.
; APPLICANT: Maiti, Pradip K.
; APPLICANT: Kaplan, Howard A.
; TITLE OF INVENTION: ANTIGEN BINDING FRAGMENTS H11, THAT
; TITLE OF INVENTION: SPECIFICALLY DETECT CANCER CELLS, NUCLEOTIDES ENCODING THE
; TITLE OF INVENTION: FRAGMENTS, AND USE THEREOF FOR THE PROPHYLAXIS AND
; TITLE OF INVENTION: DETECTION OF CANCERS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster LLP
; STREET: 755 Page Mill Road

; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/862,124
; FILING DATE: 22-MAY-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Lehnhardt, Susan K.
; REGISTRATION NUMBER: 33,943
; REFERENCE/DOCKET NUMBER: 31608-20001.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 813-5600
; TELEFAX: (650) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 179 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-862-124-2

Query Match 61.2%; Score 52; DB 3; Length 179;
Best Local Similarity 70.6%; Pred. No. 0.3;
Matches 12; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 VISENGRTINYADSVKG 17
||| : |||||
DB 111 VISYDGTXYADSVKG 127

RESULT 23
US-08-862-124-17
; Sequence 17, Application US/08862124
; Patent No. 6207153
; GENERAL INFORMATION:
; APPLICANT: Dan, Michael D.
; APPLICANT: Maiti, Pradip K.
; APPLICANT: Kaplan, Howard A.
; TITLE OF INVENTION: ANTIGEN BINDING FRAGMENTS H11, THAT
; TITLE OF INVENTION: SPECIFICALLY DETECT CANCER CELLS, NUCLEOTIDES ENCODING THE
; TITLE OF INVENTION: FRAGMENTS, AND USE THEREOF FOR THE PROPHYLAXIS AND
; TITLE OF INVENTION: DETECTION OF CANCERS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster LLP
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/862,124
; FILING DATE: 22-MAY-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Lehnhardt, Susan K.
; REGISTRATION NUMBER: 33,943
; REFERENCE/DOCKET NUMBER: 31608-20001.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 813-5600

TELEFAX: (650) 494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 287 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-862-124-17

Query Match 61.2%; Score 52; DB 3; Length 287;
Best Local Similarity 70.6%; Pred. No. 0.52;
Matches 12; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 VISENGRTINYADSVKG 17
|||:|||||
Db 192 VISYDGTKYADSVKG 208

RESULT 24
US-08-862-124-14
Sequence 14, Application US/08862124
Patent No. 6207153

GENERAL INFORMATION:
APPLICANT: Dan, Michael D.
APPLICANT: Maiti, Pradip K.
APPLICANT: Kaplan, Howard A.
TITLE OF INVENTION: ANTIGEN BINDING FRAGMENTS H11, THAT
SPECIFICALLY DETECT CANCER CELLS, NUCLEOTIDES ENCODING THE
TITLE OF INVENTION: FRAGMENTS, AND USE THEREOF FOR THE PROPHYLAXIS AND
TITLE OF INVENTION: DETECTION OF CANCERS

NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSER: Morrison & Foerster LLP
STREET: 755 Page Mill Road
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1018

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: 22-MAY-1997
APPLICATION NUMBER: US/08/862,124
CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:
NAME: Lehnhardt, Susan K.
REGISTRATION NUMBER: 33,943
REFERENCE/DOCKET NUMBER: 31608-20001.20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 813-5600
TELEFAX: (650) 494-0792
TELEX: 706141

INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 304 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-862-124-14

Query Match 61.2%; Score 52; DB 3; Length 304;
Best Local Similarity 70.6%; Pred. No. 0.55;
Matches 12; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 VISENGRTINYADSVKG 17
|||:|||||
Db 209 VISYDGTKYADSVKG 225

RESULT 25

US-08-211-202-116
Sequence 116, Application US/08211202
Patent No. 5565332

GENERAL INFORMATION:
APPLICANT: HOOGENBOOM, Hendricus Renerus Jacobus Matteus
APPLICANT: BAIER, Michael
APPLICANT: JESPEERS, Laurent Stephane Anne Therese
APPLICANT: WINTER, Gregory Paul
TITLE OF INVENTION: Production of chimeric antibodies - a
TITLE OF INVENTION: combinatorial approach
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESS:

ADDRESSEE: David W. Clough, Marshall O'Toole Gerstein Murray &
ADDRESSEE: Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606-6402

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/211,202
FILING DATE: 23-SEP-1992
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9120252.3
FILING DATE: 23-SEP-1991

PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9120377.8
FILING DATE: 25-SEP-1991

PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9206318.9
FILING DATE: 24-MAR-1992

PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9206372.6
FILING DATE: 24-MAR-1992

PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB92/00883
FILING DATE: 15-MAY-1992

ATTORNEY/AGENT INFORMATION:
NAME: David W. Clough

REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 28111/31960
TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 116:
SEQUENCE CHARACTERISTICS:
LENGTH: 98 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
US-08-211-202-116

Query Match 60.0%; Score 51; DB 1; Length 98;
Best Local Similarity 68.8%; Pred. No. 0.23;
Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVKG 17
|||:|||||
Db 51 ISSSSSTIYADSVKG 66

Search completed: September 24, 2004, 02:20:03
Job time : 30.8679 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 01:54:43 ; Search time 98.4717 Seconds
(without alignments)
55.513 Million cell updates/sec

Title: US-10-088-639A-2_COPY_177_193

Perfect score: 85

Sequence: 1 VISENGRTINADSVKG 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1349238 seqs, 32158718 residues

Total number of hits satisfying chosen parameters: 1349238

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pap:*
- 2: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pap:*
- 3: /cgn2_6/ptodata/2/pubpaa/US05_NEW_PUB.pap:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pap:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pap:*
- 6: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pap:*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pap:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pap:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pap:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pap:*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pap:*
- 12: /cgn2_6/ptodata/2/pubpaa/US09D_PUBCOMB.pap:*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pap:*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pap:*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pap:*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pap:*
- 17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pap:*
- 18: /cgn2_6/ptodata/2/pubpaa/US10F_PUBCOMB.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	65	76.5	17	14	US-10-031-874A-172
2	61	71.8	248	10	US-09-880-748-1648
3	61	71.8	248	12	US-10-293-418-1648
4	59	69.4	129	10	US-09-848-798-143
5	58	68.2	261	10	US-09-880-748-1503
6	58	68.2	261	12	US-10-293-418-1503
7	57	67.1	98	12	US-10-453-698-57
8	57	67.1	98	12	US-10-029-926B-63
9	57	67.1	98	12	US-10-029-926B-79
10	57	67.1	98	13	US-10-066-895-4
11	57	67.1	98	14	US-10-194-975-17
12	57	67.1	98	15	US-10-308-817-57
13	57	67.1	98	15	US-10-032-037B-63
14	57	67.1	98	15	US-10-032-037B-79
15	57	67.1	98	15	US-10-029-988B-63

16	57	67.1	98	15	US-10-029-988B-79	Sequence 79, Appl
17	57	67.1	98	15	US-10-032-423A-63	Sequence 63, Appl
18	57	67.1	98	15	US-10-032-423A-79	Sequence 79, Appl
19	57	67.1	98	16	US-10-038-591-30	Sequence 30, Appl
20	57	67.1	98	16	US-10-379-392-16	Sequence 16, Appl
21	57	67.1	109	15	US-10-309-764-23	Sequence 23, Appl
22	57	67.1	122	14	US-10-269-805-7	Sequence 7, Appl
23	57	67.1	122	14	US-10-269-805-11	Sequence 11, Appl
24	57	67.1	122	14	US-10-269-805-53	Sequence 53, Appl
25	57	67.1	131	10	US-09-988-115A-53	Sequence 53, Appl
26	57	67.1	236	10	US-09-880-748-2010	Sequence 2010, Ap
27	57	67.1	236	12	US-10-293-418-2010	Sequence 2010, Ap
28	57	67.1	237	10	US-09-880-748-2003	Sequence 2003, Ap
29	57	67.1	237	10	US-09-880-748-2005	Sequence 2005, Ap
30	57	67.1	237	10	US-09-880-748-2006	Sequence 2006, Ap
31	57	67.1	237	10	US-09-880-748-2007	Sequence 2007, Ap
32	57	67.1	237	10	US-09-880-748-2017	Sequence 2017, Ap
33	57	67.1	237	10	US-09-880-748-2019	Sequence 2019, Ap
34	57	67.1	237	10	US-09-880-748-2020	Sequence 2020, Ap
35	57	67.1	237	10	US-09-880-748-2027	Sequence 2027, Ap
36	57	67.1	237	10	US-09-880-748-2028	Sequence 2028, Ap
37	57	67.1	237	10	US-09-880-748-2036	Sequence 2036, Ap
38	57	67.1	237	10	US-09-880-748-2040	Sequence 2040, Ap
39	57	67.1	237	10	US-09-880-748-2043	Sequence 2043, Ap
40	57	67.1	237	10	US-09-880-748-2104	Sequence 2104, Ap
41	57	67.1	237	10	US-09-880-748-2110	Sequence 2110, Ap
42	57	67.1	237	10	US-09-880-748-2111	Sequence 2111, Ap
43	57	67.1	237	10	US-09-880-748-2112	Sequence 2112, Ap
44	57	67.1	237	10	US-09-880-748-2114	Sequence 2114, Ap
45	57	67.1	237	10	US-09-880-748-2115	Sequence 2115, Ap
46	57	67.1	237	12	US-10-293-418-2118	Sequence 2118, Ap
47	57	67.1	237	12	US-10-293-418-2003	Sequence 2003, Ap
48	57	67.1	237	12	US-10-293-418-2005	Sequence 2005, Ap
49	57	67.1	237	12	US-10-293-418-2006	Sequence 2006, Ap
50	57	67.1	237	12	US-10-293-418-2017	Sequence 2017, Ap
51	57	67.1	237	12	US-10-293-418-2019	Sequence 2019, Ap
52	57	67.1	237	12	US-10-293-418-2020	Sequence 2020, Ap
53	57	67.1	237	12	US-10-293-418-2027	Sequence 2027, Ap
54	57	67.1	237	12	US-10-293-418-2028	Sequence 2028, Ap
55	57	67.1	237	12	US-10-293-418-2036	Sequence 2036, Ap
56	57	67.1	237	12	US-10-293-418-2040	Sequence 2040, Ap
57	57	67.1	237	12	US-10-293-418-2043	Sequence 2043, Ap
58	57	67.1	237	12	US-10-293-418-2104	Sequence 2104, Ap
59	57	67.1	237	12	US-10-293-418-2110	Sequence 2110, Ap
60	57	67.1	237	12	US-10-293-418-2111	Sequence 2111, Ap
61	57	67.1	237	12	US-10-293-418-2112	Sequence 2112, Ap
62	57	67.1	237	12	US-10-293-418-2114	Sequence 2114, Ap
63	57	67.1	237	12	US-10-293-418-2115	Sequence 2115, Ap
64	57	67.1	237	12	US-10-293-418-2118	Sequence 2118, Ap
65	57	67.1	238	12	US-09-880-748-2024	Sequence 2024, Ap
66	57	67.1	239	10	US-09-880-748-2034	Sequence 2034, Ap
67	57	67.1	239	12	US-10-293-418-2034	Sequence 2034, Ap
68	57	67.1	240	10	US-09-880-748-2007	Sequence 2007, Ap
69	57	67.1	240	10	US-09-880-748-2016	Sequence 2016, Ap
70	57	67.1	240	10	US-09-880-748-2025	Sequence 2025, Ap
71	57	67.1	240	10	US-09-880-748-2029	Sequence 2029, Ap
72	57	67.1	240	10	US-09-880-748-2030	Sequence 2030, Ap
73	57	67.1	240	10	US-09-880-748-2041	Sequence 2041, Ap
74	57	67.1	240	10	US-09-880-748-2044	Sequence 2044, Ap
75	57	67.1	240	10	US-09-880-748-2045	Sequence 2045, Ap
76	57	67.1	240	10	US-09-880-748-2048	Sequence 2048, Ap
77	57	67.1	240	10	US-09-880-748-2105	Sequence 2105, Ap
78	57	67.1	240	10	US-09-880-748-2108	Sequence 2108, Ap
79	57	67.1	240	10	US-09-880-748-2113	Sequence 2113, Ap
80	57	67.1	240	10	US-09-880-748-2117	Sequence 2117, Ap
81	57	67.1	240	12	US-10-293-418-2117	Sequence 2117, Ap
82	57	67.1	240	12	US-10-293-418-2016	Sequence 2016, Ap
83	57	67.1	240	12	US-10-293-418-2025	Sequence 2025, Ap
84	57	67.1	240	12	US-10-293-418-2029	Sequence 2029, Ap
85	57	67.1	240	12	US-10-293-418-2030	Sequence 2030, Ap
86	57	67.1	240	12	US-10-293-418-2041	Sequence 2041, Ap
87	57	67.1	240	12	US-10-293-418-2044	Sequence 2044, Ap
88	57	67.1	240	12	US-10-293-418-2045	Sequence 2045, Ap

; PRIOR APPLICATION NUMBER: 60/331,469
 ; PRIOR FILING DATE: 2001-11-16
 ; PRIOR APPLICATION NUMBER: 60/340,817
 ; PRIOR FILING DATE: 2001-12-19
 ; PRIOR APPLICATION NUMBER: 09/880,748
 ; PRIOR FILING DATE: 2001-06-15
 ; PRIOR APPLICATION NUMBER: 60/293,499
 ; PRIOR FILING DATE: 2001-05-25
 ; PRIOR APPLICATION NUMBER: 60/277,379
 ; PRIOR FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: 60/276,248
 ; PRIOR FILING DATE: 2001-03-16
 ; PRIOR APPLICATION NUMBER: 60/240,816
 ; PRIOR FILING DATE: 2000-10-17
 ; PRIOR APPLICATION NUMBER: 60/212,210
 ; PRIOR FILING DATE: 2000-06-16
 ; NUMBER OF SEQ ID NOS: 3247
 ; SEQ ID NO 1648
 ; LENGTH: 248
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-293-418-1648

Query Match 71.8%; Score 61; DB 12; Length 248;
 Best Local Similarity 64.7%; Pred. No. 0.045;
 Matches 11; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 QY 1 VISENGRTINYADSVKG 17
 |||:|:|:|:|
 Db 50 VISDDGNVNYEDSVKG 66

RESULT 4
 US-09-848-798-143
 ; Sequence 143, Application US/09848798
 ; Publication No. US20030040605A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Siegel, Donald L.
 ; TITLE OF INVENTION: RH(D)-BINDING PROTEINS AND MAGNETICALLY ACTIVATED CELL
 ; TITLE OF INVENTION: SORTING METHOD FOR PRODUCTION THEREOF
 ; FILE REFERENCE: 09596-42U2
 ; CURRENT APPLICATION NUMBER: US/09/848,798
 ; PRIOR FILING DATE: 2001-05-04
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/240,274
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1999-01-29
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/028,550
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1996-10-11
 ; NUMBER OF SEQ ID NOS: 224
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 143
 ; LENGTH: 129
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: anti-Rh(D) antibody clone SH20
 US-09-848-798-143

Query Match 69.4%; Score 59; DB 10; Length 129;
 Best Local Similarity 76.5%; Pred. No. 0.046;
 Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 1 VISENGRTINYADSVKG 17
 |||:|:|:|:|
 Db 51 VISYDGTIYADSVKG 67

RESULT 5
 US-09-880-748-1503
 ; Sequence 1503, Application US/09880748
 ; Publication No. US2003005937A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ruben et al.
 ; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys

; FILE REFERENCE: PF523
 ; CURRENT APPLICATION NUMBER: US/09/880,748
 ; CURRENT FILING DATE: 2001-06-15
 ; PRIOR APPLICATION NUMBER: 60/212,210
 ; PRIOR FILING DATE: 2000-06-15
 ; PRIOR APPLICATION NUMBER: 60/240,816
 ; PRIOR FILING DATE: 2000-10-17
 ; PRIOR APPLICATION NUMBER: 60/276,248
 ; PRIOR FILING DATE: 2001-03-16
 ; PRIOR APPLICATION NUMBER: 60/277,379
 ; PRIOR FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: 60/293,499
 ; PRIOR FILING DATE: 2001-05-25
 ; NUMBER OF SEQ ID NOS: 3239
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 1503
 ; LENGTH: 261
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-880-748-1503

Query Match 68.2%; Score 58; DB 10; Length 261;
 Best Local Similarity 76.5%; Pred. No. 0.15;
 Matches 13; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 VISENGRTINYADSVKG 17
 |||:|:|:|:|
 Db 52 VISYNGRTKYLDVSKG 68

RESULT 6
 US-10-293-418-1503
 ; Sequence 1503, Application US/10293418
 ; Publication No. US20030223996A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ruben et al.
 ; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
 ; FILE REFERENCE: PF523P2
 ; CURRENT APPLICATION NUMBER: US/10/293,418
 ; CURRENT FILING DATE: 2002-11-27
 ; PRIOR APPLICATION NUMBER: 60/331,469
 ; PRIOR FILING DATE: 2001-11-16
 ; PRIOR APPLICATION NUMBER: 60/340,817
 ; PRIOR FILING DATE: 2001-12-19
 ; PRIOR APPLICATION NUMBER: 09/880,748
 ; PRIOR FILING DATE: 2001-06-15
 ; PRIOR APPLICATION NUMBER: 60/293,499
 ; PRIOR FILING DATE: 2001-05-25
 ; PRIOR APPLICATION NUMBER: 60/277,379
 ; PRIOR FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: 60/276,248
 ; PRIOR FILING DATE: 2001-03-16
 ; PRIOR APPLICATION NUMBER: 60/240,816
 ; PRIOR FILING DATE: 2000-10-17
 ; PRIOR APPLICATION NUMBER: 60/212,210
 ; PRIOR FILING DATE: 2000-06-16
 ; NUMBER OF SEQ ID NOS: 3247
 ; SEQ ID NO 1503
 ; LENGTH: 261
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-293-418-1503

Query Match 68.2%; Score 58; DB 12; Length 261;
 Best Local Similarity 76.5%; Pred. No. 0.15;
 Matches 13; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 VISENGRTINYADSVKG 17
 |||:|:|:|:|
 Db 52 VISYNGRTKYLDVSKG 68

RESULT 7

```
US-10-453-698-57
; Sequence 57, Application US/10453698
; Publication No. US20040038308A1
; GENERAL INFORMATION:
; APPLICANT: Rother, Russell
; TITLE OF INVENTION: HYBRID ANTIBODIES
; FILE REFERENCE: 82 CIP (1087-37 CIP)
; CURRENT APPLICATION NUMBER: US/10/453,698
; CURRENT FILING DATE: 2003-08-03
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 57
; LENGTH: 98
; TYPE: PRT
; ORGANISM: human
US-10-453-698-57

Query Match      67.1%; Score 57; DB 12; Length 98;
Best Local Similarity 75.0%; Pred. No. 0.073;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVKG 17
   ||:|||||
Db 51 ISSSGSTIYADSVKG 66

RESULT 8
US-10-029-926B-63
; Sequence 63, Application US/10029926B
; Publication No. US20040073011A1
; GENERAL INFORMATION:
; APPLICANT: HAGAY, et al.
; TITLE OF INVENTION: SPECIFIC HUMAN ANTIBODIES FOR SELECTIVE CANCER THERAPY
; FILE REFERENCE: 10793/50
; CURRENT APPLICATION NUMBER: US/10/029,926B
; CURRENT FILING DATE: 2001-12-31
; PRIOR APPLICATION NUMBER: 60/258,948
; PRIOR FILING DATE: 12/29/2000
; NUMBER OF SEQ ID NOS: 203
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 63
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-029-926B-63

Query Match      67.1%; Score 57; DB 12; Length 98;
Best Local Similarity 75.0%; Pred. No. 0.073;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVKG 17
   ||:|||||
Db 51 ISSSGSTIYADSVKG 66

RESULT 9
US-10-029-926B-79
; Sequence 79, Application US/10029926B
; Publication No. US20040073011A1
; GENERAL INFORMATION:
; APPLICANT: HAGAY, et al.
; TITLE OF INVENTION: SPECIFIC HUMAN ANTIBODIES FOR SELECTIVE CANCER THERAPY
; FILE REFERENCE: 10793/50
; CURRENT APPLICATION NUMBER: US/10/029,926B
; CURRENT FILING DATE: 2001-12-31
; PRIOR APPLICATION NUMBER: 60/258,948
; PRIOR FILING DATE: 12/29/2000
; NUMBER OF SEQ ID NOS: 203
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 79
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-029-926B-79

Query Match      67.1%; Score 57; DB 12; Length 98;
Best Local Similarity 75.0%; Pred. No. 0.073;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVKG 17
   ||:|||||
Db 51 ISSSGSTIYADSVKG 66

US-10-029-926B-79
Query Match      67.1%; Score 57; DB 12; Length 98;
Best Local Similarity 75.0%; Pred. No. 0.073;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVKG 17
   ||:|||||
Db 51 ISSSGSTIYADSVKG 66

RESULT 10
US-10-066-895-4
; Sequence 4, Application US/10066895
; Publication No. US20020141990A1
; GENERAL INFORMATION:
; APPLICANT: Deen, Keith C.
;             Dillon, Susan B.
;             Porter, Terence C.
;             Sweet, Raymond A.
; TITLE OF INVENTION: Human Monoclonal Antibodies
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19046
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/066,895
; FILING DATE: 04-Feb-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/297,344
; FILING DATE: 1999-JUN-09
; APPLICATION NUMBER: 60/030,149
; FILING DATE: 01-NOV-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Geiger, Kathleen
; REGISTRATION NUMBER: 35,880
; REFERENCE/DOCKET NUMBER: P50504
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5368
; TELEFAX: 610-270-5090
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 98 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-066-895-4

Query Match      67.1%; Score 57; DB 13; Length 98;
Best Local Similarity 75.0%; Pred. No. 0.073;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVKG 17
   ||:|||||
Db 51 ISSSGSTIYADSVKG 66

RESULT 11
US-10-194-975-17
; Sequence 17, Application US/10194975
```

Publication No. US20030039649A1
; GENERAL INFORMATION:
; APPLICANT: Foote, Jefferson
; TITLE OF INVENTION: Super Humanized Antibodies
; FILE REFERENCE: 501231.01
; CURRENT APPLICATION NUMBER: US/10/194,975
; CURRENT FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: US 60/305,111
; PRIOR FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-194-975-17

Query Match 67.1%; Score 57; DB 14; Length 98;
Best Local Similarity 75.0%; Pred. No. 0.073;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVK 17
||:|||||
Db 51 ISSSGSTIYADSVK 66

RESULT 12

US-10-308-817-57
; Sequence 57, Application US/10308817
; Publication No. US20030219861A1
; GENERAL INFORMATION:
; APPLICANT: Rother, Russell
; APPLICANT: Wu, Dayang
; TITLE OF INVENTION: HYBRID ANTIBODIES
; FILE REFERENCE: 1087-37
; CURRENT APPLICATION NUMBER: US/10/308,817
; CURRENT FILING DATE: 2002-12-03
; NUMBER OF SEQ ID NOS: 195
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 57
; LENGTH: 98
; TYPE: PRT
; ORGANISM: human
US-10-308-817-57

Query Match 67.1%; Score 57; DB 15; Length 98;
Best Local Similarity 75.0%; Pred. No. 0.073;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVK 17
||:|||||
Db 51 ISSSGSTIYADSVK 66

RESULT 13

US-10-032-037B-63
; Sequence 63, Application US/10032037B
; Publication No. US20040001822A1
; GENERAL INFORMATION:
; APPLICANT: Bio-Technology General Corp.
; TITLE OF INVENTION: Y17-ISOLATED MOLECULES COMPRISING EPITOPES CONTAINING SULFATED
; FILE REFERENCE: 10793/44
; CURRENT APPLICATION NUMBER: US/10/032,037B
; CURRENT FILING DATE: 2001-12-31
; PRIOR APPLICATION NUMBER: 60/258,948
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 204
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 63
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Homo sapiens

US-10-032-037B-63

Query Match 67.1%; Score 57; DB 15; Length 98;
Best Local Similarity 75.0%; Pred. No. 0.073;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVK 17
||:|||||
Db 51 ISSSGSTIYADSVK 66

RESULT 14

US-10-032-037B-79
; Sequence 79, Application US/10032037B
; Publication No. US20040001822A1
; GENERAL INFORMATION:
; APPLICANT: Bio-Technology General Corp.
; TITLE OF INVENTION: Y17-ISOLATED MOLECULES COMPRISING EPITOPES CONTAINING SULFATED
; FILE REFERENCE: 10793/44
; CURRENT APPLICATION NUMBER: US/10/032,037B
; CURRENT FILING DATE: 2001-12-31
; PRIOR APPLICATION NUMBER: 60/258,948
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 204
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 79
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-032-037B-79

Query Match 67.1%; Score 57; DB 15; Length 98;
Best Local Similarity 75.0%; Pred. No. 0.073;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVK 17
||:|||||
Db 51 ISSSGSTIYADSVK 66

RESULT 15

US-10-029-988B-63
; Sequence 63, Application US/10029988B
; Publication No. US20040001839A1
; GENERAL INFORMATION:
; APPLICANT: Bio-Technology General Corp.
; TITLE OF INVENTION: Y17-ISOLATED MOLECULES COMPRISING EPITOPES CONTAINING SULFATED
; FILE REFERENCE: 10793/46
; CURRENT APPLICATION NUMBER: US/10/029,988B
; CURRENT FILING DATE: 2001-12-31
; PRIOR APPLICATION NUMBER: 60/258,948
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 204
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 63
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-029-988B-63

Query Match 67.1%; Score 57; DB 15; Length 98;
Best Local Similarity 75.0%; Pred. No. 0.073;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVK 17
||:|||||
Db 51 ISSSGSTIYADSVK 66

RESULT 16

US-10-029-988B-79

```
; Sequence 79, Application US/10029988B
; Publication No. US20040001839A1
; GENERAL INFORMATION:
; APPLICANT: Bio-Technology General Corp.
; TITLE OF INVENTION: Y17-ISOLATED MOLECULES COMPRISING EPITOPES CONTAINING SULFATED
; MOIETIES, ANTIBODIES TO SUCH EPITOPES, AND USES THEREOF
; FILE REFERENCE: 10793/46
; CURRENT APPLICATION NUMBER: US/10/029,988B
; CURRENT FILING DATE: 2001-12-31
; PRIOR APPLICATION NUMBER: 60/258,948
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 204
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 79
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-029-988B-79

Query Match      67.1%; Score 57; DB 15; Length 98;
Best Local Similarity 75.0%; Pred. No. 0.073;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      2 ISENGRTINYADSVKG 17
DB      51 ISSSGSTIYYADSVKG 66

RESULT 17
US-10-032-423A-63
; Sequence 63, Application US/10032423A
; Publication No. US20040002450A1
; GENERAL INFORMATION:
; APPLICANT: Bio-Technology General Corp.
; TITLE OF INVENTION: Y17-ISOLATED MOLECULES COMPRISING EPITOPES CONTAINING SULFATED
; MOIETIES, ANTIBODIES TO SUCH EPITOPES, AND USES THEREOF
; FILE REFERENCE: 10793/45
; CURRENT APPLICATION NUMBER: US/10/032,423A
; CURRENT FILING DATE: 2001-12-31
; PRIOR APPLICATION NUMBER: 60/258,948
; PRIOR FILING DATE: 12/29/2000
; NUMBER OF SEQ ID NOS: 204
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 63
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-032-423A-63

Query Match      67.1%; Score 57; DB 15; Length 98;
Best Local Similarity 75.0%; Pred. No. 0.073;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      2 ISENGRTINYADSVKG 17
DB      51 ISSSGSTIYYADSVKG 66

RESULT 18
US-10-032-423A-79
; Sequence 79, Application US/10032423A
; Publication No. US20040002450A1
; GENERAL INFORMATION:
; APPLICANT: Bio-Technology General Corp.
; TITLE OF INVENTION: Y17-ISOLATED MOLECULES COMPRISING EPITOPES CONTAINING SULFATED
; MOIETIES, ANTIBODIES TO SUCH EPITOPES, AND USES THEREOF
; FILE REFERENCE: 10793/45
; CURRENT APPLICATION NUMBER: US/10/032,423A
; CURRENT FILING DATE: 2001-12-31
; PRIOR APPLICATION NUMBER: 60/258,948
; PRIOR FILING DATE: 12/29/2000
; NUMBER OF SEQ ID NOS: 204
; SOFTWARE: FastSEQ for Windows Version 3.0
```

```
; SEQ ID NO 79
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-032-423A-79

Query Match      67.1%; Score 57; DB 15; Length 98;
Best Local Similarity 75.0%; Pred. No. 0.073;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      2 ISENGRTINYADSVKG 17
DB      51 ISSSGSTIYYADSVKG 66

RESULT 19
US-10-038-591-30
; Sequence 30, Application US/10038591
; Publication No. US20040086503A1
; GENERAL INFORMATION:
; APPLICANT: Cohen, Bruce D.
; APPLICANT: Beebe, Jean
; APPLICANT: Miller, Penelope E.
; APPLICANT: Moyer, James D.
; APPLICANT: Corvalan, Jose R.
; APPLICANT: Gallo, Michael
; TITLE OF INVENTION: ANTIBODIES TO INSULIN-LIKE GROWTH FACTOR I RECEPTOR
; FILE REFERENCE: ABX-PF2
; CURRENT APPLICATION NUMBER: US/10/038,591
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 60/259,927
; PRIOR FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-038-591-30

Query Match      67.1%; Score 57; DB 16; Length 98;
Best Local Similarity 75.0%; Pred. No. 0.073;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      2 ISENGRTINYADSVKG 17
DB      51 ISSSGSTIYYADSVKG 66

RESULT 20
US-10-379-392-16
; Sequence 16, Application US/10379392
; Publication No. US20040110226A1
; GENERAL INFORMATION:
; APPLICANT: Lazar, Gregory Alan
; APPLICANT: Desjarlais, John Rudolf
; APPLICANT: Marshall, Shannon Alicia
; APPLICANT: Dahiyat, Bassil I.
; TITLE OF INVENTION: ANTIBODY OPTIMIZATION
; FILE REFERENCE: A-71386-3 463077-236
; CURRENT APPLICATION NUMBER: US/10/379,392
; CURRENT FILING DATE: 2003-03-03
; PRIOR APPLICATION NUMBER: US 60/360,843
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 60/384,197
; PRIOR FILING DATE: 2002-05-29
; NUMBER OF SEQ ID NOS: 184
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 16
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-379-392-16
```

```

Query Match      67.1%; Score 57; DB 16; Length 98;
Best Local Similarity 75.0%; Pred. No. 0.073;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVKG 17
DB 51 ISSSGSTIYYADSVKG 66

RESULT 21
US-10-309-764-23
; Sequence 23, Application US/10309764
; Publication No. US20030322009A1
; GENERAL INFORMATION:
; APPLICANT: Foltz, Ian
; APPLICANT: Babcock, John
; APPLICANT: Palathumpat, Raju
; APPLICANT: Yang, Xiao-dong
; APPLICANT: King, Chadwick T.
; TITLE OF INVENTION: ANTI-CDR45RB ANTIBODIES FOR USE IN
; TITLE OF INVENTION: TREATING AUTOIMMUNE DISEASE AND TRANSPLANT REJECTION
; FILE REFERENCE: ABGENIX 029A
; CURRENT APPLICATION NUMBER: US/10/309,764
; CURRENT FILING DATE: 2002-12-02
; PRIOR APPLICATION NUMBER: 60/337,276
; PRIOR FILING DATE: 2001-12-03
; NUMBER OF SEQ ID NOS: 147
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 109
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-309-764-23

Query Match      67.1%; Score 57; DB 15; Length 109;
Best Local Similarity 75.0%; Pred. No. 0.082;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVKG 17
DB 51 ISSSGSTIYYADSVKG 66

RESULT 22
US-10-269-805-7
; Sequence 7, Application US/10269805
; Publication No. US20030124129A1
; GENERAL INFORMATION:
; APPLICANT: OLINER, JONATHAN D.
; TITLE OF INVENTION: ANGIOPOIETIN-2 SPECIFIC BINDING AGENTS
; FILE REFERENCE: A-722
; CURRENT APPLICATION NUMBER: US/10/269,805
; CURRENT FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: US 60/328,604
; PRIOR FILING DATE: 2001-10-11
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 122
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-269-805-7

Query Match      67.1%; Score 57; DB 14; Length 122;
Best Local Similarity 75.0%; Pred. No. 0.094;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVKG 17
DB 51 ISSSGSTIYYADSVKG 66

RESULT 25
US-09-988-115A-53
; Sequence 53, Application US/09988115A
; Publication No. US20030037347A1
; GENERAL INFORMATION:
; APPLICANT: Robl, James M.
; APPLICANT: Goldsby, Richard A.
; APPLICANT: Ferguson, Stacy B.
; APPLICANT: Kuroiwa, Yoshima
; APPLICANT: Tomizuka, Kazuma
; APPLICANT: Ishida, Isao
; TITLE OF INVENTION: Expression of Xenogenous (Human)
; TITLE OF INVENTION: Immunoglobulins in Cloned, Transgenic Ungulates
; FILE REFERENCE: 50195/008003
; CURRENT APPLICATION NUMBER: US/09/988,115A

```

```

RESULT 23
US-10-269-805-11
; Sequence 11, Application US/10269805
; Publication No. US20030124129A1
; GENERAL INFORMATION:
; APPLICANT: OLINER, JONATHAN D.
; TITLE OF INVENTION: ANGIOPOIETIN-2 SPECIFIC BINDING AGENTS
; FILE REFERENCE: A-722
; CURRENT APPLICATION NUMBER: US/10/269,805
; CURRENT FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: US 60/328,604
; PRIOR FILING DATE: 2001-10-11
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 122
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-269-805-11

Query Match      67.1%; Score 57; DB 14; Length 122;
Best Local Similarity 75.0%; Pred. No. 0.094;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVKG 17
DB 51 ISSSGSTIYYADSVKG 66

RESULT 24
US-10-269-805-53
; Sequence 53, Application US/10269805
; Publication No. US20030124129A1
; GENERAL INFORMATION:
; APPLICANT: OLINER, JONATHAN D.
; TITLE OF INVENTION: ANGIOPOIETIN-2 SPECIFIC BINDING AGENTS
; FILE REFERENCE: A-722
; CURRENT APPLICATION NUMBER: US/10/269,805
; CURRENT FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: US 60/328,604
; PRIOR FILING DATE: 2001-10-11
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 53
; LENGTH: 122
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-269-805-53

Query Match      67.1%; Score 57; DB 14; Length 122;
Best Local Similarity 75.0%; Pred. No. 0.094;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVKG 17
DB 51 ISSSGSTIYYADSVKG 66

RESULT 25
US-09-988-115A-53
; Sequence 53, Application US/09988115A
; Publication No. US20030037347A1
; GENERAL INFORMATION:
; APPLICANT: Robl, James M.
; APPLICANT: Goldsby, Richard A.
; APPLICANT: Ferguson, Stacy B.
; APPLICANT: Kuroiwa, Yoshima
; APPLICANT: Tomizuka, Kazuma
; APPLICANT: Ishida, Isao
; TITLE OF INVENTION: Expression of Xenogenous (Human)
; TITLE OF INVENTION: Immunoglobulins in Cloned, Transgenic Ungulates
; FILE REFERENCE: 50195/008003
; CURRENT APPLICATION NUMBER: US/09/988,115A

```

```

; CURRENT FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,625
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: US 60/256,458
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: US 09/714,185
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: US 60/166,410
; PRIOR FILING DATE: 1999-11-19
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 53
; LENGTH: 131
; TYPE: PRT
; ORGANISM: Bovine
; US-09-988-115A-53

```

```

Query Match      67.1%; Score 57; DB 10; Length 131;
Best Local Similarity 75.0%; Pred. No. 0.1;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      2 ISENGRTINYADSVKG 17
      ||:|||||
Db      43 ISSGSGTIYYADSVKG 58

```

Search completed: September 24, 2004, 02:25:14
Job time : 99.4717 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 01:33:11 ; Search time 25.3396 Seconds
(without alignments)
64.534 Million cell updates/sec

Title: US-10-088-639A-2_COPY_177_193
Perfect score: 85
Sequence: 1 VISENGRTINYADSVKG 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database : PIR 78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	58	68.2	115	2 S57445	Ig heavy chain V-J
2	57	67.1	97	2 S24252	Ig heavy chain V r
3	57	67.1	98	2 S26930	Ig heavy chain V r
4	57	67.1	98	2 S26891	Ig heavy chain V r
5	57	67.1	110	2 PH1652	Ig heavy chain V r
6	57	67.1	114	2 S31120	Ig heavy chain - h
7	57	67.1	118	2 S31105	Ig heavy chain (su
8	57	67.1	128	2 S26790	Ig heavy chain V r
9	57	67.1	130	2 I37783	Ig variable region
10	57	67.1	132	2 I47191	Ig heavy chain var
11	56	65.9	97	2 PH0875	Ig heavy chain V r
12	53	62.4	120	2 S36273	Ig heavy chain V r
13	52	61.2	108	2 S24251	Ig heavy chain V r
14	52	61.2	115	2 S36267	Ig heavy chain V r
15	52	61.2	123	2 S30532	Ig heavy chain V r
16	51	60.0	98	2 S26894	Ig heavy chain V r
17	51	60.0	110	2 S36282	Ig heavy chain V r
18	51	60.0	114	1 M3HUME	Ig heavy chain V-I
19	51	60.0	117	2 S21980	Ig heavy chain V-I
20	51	60.0	121	2 PH1661	Ig heavy chain V r
21	51	60.0	123	2 S26794	Ig heavy chain V r
22	51	60.0	125	2 S30531	Ig heavy chain V r
23	51	60.0	143	2 S23624	Ig heavy chain V r
24	50	58.8	98	2 PH0874	Ig heavy chain V r
25	50	58.8	98	2 S26927	Ig heavy chain V r
26	50	58.8	98	2 S26934	Ig heavy chain V r
27	50	58.8	113	2 S57441	Ig heavy chain V-J
28	50	58.8	121	2 S31104	Ig heavy chain (su
29	50	58.8	128	2 S31595	Ig heavy chain V r

30	50	58.8	133	2 I47195	Ig heavy chain var
31	50	58.8	151	2 A60943	Ig heavy chain pre
32	50	58.8	191	2 JL0048	Ig heavy chain V r
33	49	57.6	95	2 S20777	Ig heavy chain V r
34	49	57.6	120	1 GIHUBD	Ig heavy chain V-I
35	49	57.6	120	2 E49590	Ig heavy chain V r
36	49	57.6	121	2 B34871	Ig heavy chain V r
37	49	57.6	122	1 AlHUTR	Ig heavy chain V-I
38	49	57.6	123	2 PC4281	anti-SS-A/Ro 60K p
39	48	56.5	94	2 D25913	Ig heavy chain V r
40	48	56.5	96	2 S20781	Ig heavy chain V r
41	48	56.5	98	2 S29545	Ig heavy chain V r
42	48	56.5	98	2 S26928	Ig heavy chain V r
43	48	56.5	118	2 S00700	Ig heavy chain V r
44	48	56.5	135	2 I37778	Ig variable region
45	48	56.5	136	1 GIMS21	Ig heavy chain pre
46	48	56.5	141	2 I47178	Ig heavy chain var
47	47	55.3	94	2 PL0120	Ig heavy chain V-I
48	47	55.3	97	2 S44115	Ig heavy chain V r
49	47	55.3	98	2 PL0116	Ig heavy chain V-I
50	47	55.3	98	2 S26896	Ig heavy chain V r
51	47	55.3	98	2 S29546	Ig heavy chain V r
52	47	55.3	102	2 PH1017	Ig heavy chain V r
53	47	55.3	108	2 PH1656	Ig heavy chain V r
54	47	55.3	109	2 PH1644	Ig heavy chain V r
55	47	55.3	109	2 PH1646	Ig heavy chain V r
56	47	55.3	111	2 PH1643	Ig heavy chain V r
57	47	55.3	111	2 PH1645	Ig heavy chain V r
58	47	55.3	113	2 S38490	Ig heavy chain - h
59	47	55.3	114	2 S46390	Ig heavy chain V r
60	47	55.3	114	2 S46392	Ig heavy chain V r
61	47	55.3	116	1 M3HUGL	Ig heavy chain V-I
62	47	55.3	117	2 S36270	Ig heavy chain V r
63	47	55.3	118	2 PH0097	Ig heavy chain V r
64	47	55.3	118	2 S31677	Ig heavy chain V r
65	47	55.3	118	2 S31116	Ig heavy chain - h
66	47	55.3	119	2 F36005	Ig heavy chain V r
67	47	55.3	120	2 S31112	Ig heavy chain - h
68	47	55.3	121	2 S19666	Ig heavy chain V r
69	47	55.3	121	2 S36005	Ig heavy chain V r
70	47	55.3	122	2 S36005	Ig heavy chain V r
71	47	55.3	122	2 S31119	Ig heavy chain - h
72	47	55.3	123	2 S38493	Ig heavy chain - h
73	47	55.3	128	2 S26786	Ig heavy chain V r
74	47	55.3	130	2 PL0098	Ig heavy chain pre
75	47	55.3	132	2 S31603	Ig heavy chain V r
76	47	55.3	134	2 S31679	Ig heavy chain V r
77	47	55.3	139	2 S31674	Ig heavy chain V r
78	46	54.1	98	2 S54856	Ig heavy chain V r
79	46	54.1	100	2 PH1016	Ig heavy chain V r
80	46	54.1	110	2 PH1016	Ig heavy chain V r
81	46	54.1	110	2 S69897	Ig heavy chain V r
82	46	54.1	111	2 PL0199	anti-DNA autoantib
83	46	54.1	117	2 A28966	Ig heavy chain pre
84	46	54.1	119	2 S24526	Ig heavy chain V r
85	46	54.1	121	2 S26798	Ig heavy chain V r
86	46	54.1	122	1 M3HUM	Ig heavy chain V-I
87	46	54.1	140	2 I47204	Ig heavy chain var
88	46	54.1	145	2 S11239	Ig heavy chain V r
89	46	54.1	145	2 I47203	Ig heavy chain var
90	46	54.1	466	2 H87419	conserved hypotet
91	45	52.9	90	2 S24248	Ig heavy chain V r
92	45	52.9	98	2 S26929	Ig heavy chain V r
93	45	52.9	98	2 S26889	Ig heavy chain V r
94	45	52.9	98	2 S26933	Ig heavy chain V r
95	45	52.9	99	2 S24259	Ig heavy chain V r
96	45	52.9	100	2 S69896	Ig heavy chain V r
97	45	52.9	100	2 S24258	Ig heavy chain V r
98	45	52.9	101	2 S24257	Ig heavy chain V r
99	45	52.9	102	2 S24260	Ig heavy chain V r
100	45	52.9	104	2 S24255	Ig heavy chain V r
101	45	52.9	105	2 S38488	Ig heavy chain - h
102	45	52.9	105	2 S24249	Ig heavy chain V r

A:Reference number: S26885; MUID:93021117; PMID:1404388
A:Accession: S26891
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-98 <TOM>
A:Cross-references: EMBL:Z12358; NID:G32935; PIDN:CAA78228.1; PID:G32936
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 67.1%; Score 57; DB 2; Length 98;
Best Local Similarity 75.0%; Pred. No. 0.046; 3; Indels 0; Gaps 0;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVK 17
DB 51 ISSSGSTIYADSVK 66

RESULT 5
PH1652
Ig heavy chain V region (clone 5D4) - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 16-Aug-1996
C:Accession: PH1652
R:Hillison, J.L.; Karr, N.S.; Oppliger, I.R.; Mannik, M.; Sasso, E.H.
J. Exp. Med. 178, 331-336, 1993
A:Title: The structural basis of germline-encoded VH3 immunoglobulin binding to staphylococcus aureus protein A
A:Reference number: PH1642; MUID:93301610; PMID:8315388
A:Accession: PH1652
A:Molecule type: mRNA
A:Residues: 1-110 <HIL>
A:Experimental source: B cell
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:7-90/Domain: immunoglobulin homology <IMM>

Query Match 67.1%; Score 57; DB 2; Length 110;
Best Local Similarity 75.0%; Pred. No. 0.052; 3; Indels 0; Gaps 0;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVK 17
DB 43 ISSSGSTIYADSVK 58

RESULT 6
S31120
Ig heavy chain - human
C:Species: Homo sapiens (man)
C>Date: 02-Dec-1993 #sequence_revision 26-May-1995 #text_change 17-Mar-1999
C:Accession: S31120
R:Raaphorst, F.M.; Timmers, E.; Kenter, M.J.H.; van Tol, M.J.D.; Vossen, J.M.; Schuurman, J.
Eur. J. Immunol. 22, 247-251, 1992
A:Title: Restricted utilization of germ-line V(H)3 genes and short diverse third complement region in human IgG1 heavy chain variable region
A:Reference number: S31104; MUID:92111633; PMID:1730252
A:Accession: S31120
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-114 <RAA>
A:Cross-references: EMBL:X62972
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1991
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 67.1%; Score 57; DB 2; Length 114;
Best Local Similarity 75.0%; Pred. No. 0.054; 3; Indels 0; Gaps 0;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVK 17
DB 51 ISSSGSTIYADSVK 66

RESULT 7

S31105
Ig heavy chain (subclass IgM) - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 22-Nov-1993 #sequence_revision 26-May-1995 #text_change 23-Jul-1999
C:Accession: S31105
R:Raaphorst, F.M.; Timmers, E.; Kenter, M.J.H.; van Tol, M.J.D.; Vossen, J.M.; Schuurman, J.
Eur. J. Immunol. 22, 247-251, 1992
A:Title: Restricted utilization of germ-line V(H)3 genes and short diverse third complement region in human IgM heavy chain variable region
A:Reference number: S31104; MUID:92111633; PMID:1730252
A:Accession: S31105
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-118 <RAA>
A:Cross-references: EMBL:X63081; NID:G32648; PIDN:CAA44803.1; PID:G32649
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1991
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 67.1%; Score 57; DB 2; Length 118;
Best Local Similarity 75.0%; Pred. No. 0.056; 3; Indels 0; Gaps 0;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVK 17
DB 51 ISSSGSTIYADSVK 66

RESULT 8

S26790
Ig heavy chain V region - human
C:Species: Homo sapiens (man)
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 20-Jun-2000
C:Accession: S26790
R:Mortari, F.; Newton, J.A.; Wang, J.Y.; Schroeder Jr., H.W.
Eur. J. Immunol. 22, 241-245, 1992
A:Title: The human cord blood antibody repertoire. Frequent usage of the V(H)7 gene family
A:Reference number: S26786; MUID:92111632; PMID:1730251
A:Accession: S26790
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-128 <MOR>
A:Cross-references: EMBL:X61013; NID:G32798; PIDN:CAA43347.1; PID:G1335128
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 67.1%; Score 57; DB 2; Length 128;
Best Local Similarity 75.0%; Pred. No. 0.061; 3; Indels 0; Gaps 0;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVK 17
DB 51 ISSSGSTIYADSVK 66

RESULT 9

I37783
Ig variable region (VDJ) (clone T24-3) - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 16-Feb-1996 #sequence_revision 16-Feb-1996 #text_change 23-Jul-1999
C:Accession: I37783; S25477
R:Demaision, C.; Chastagner, P.; Theze, J.; Zouali, M.
Proc. Natl. Acad. Sci. U.S.A. 91, 514-518, 1994
A:Title: Somatic diversification in the heavy chain variable region genes expressed by human B cells
A:Reference number: A36876; MUID:94119917; PMID:8290556
A:Accession: I37783
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-130 <RES>

A:Cross-references: EMBL:X67907; NID:G33584; PIDN:CAA48105.1; PID:G33585
C:Superfamily: immunoglobulin V region; immunoglobulin homology
F:22-105/Domain: immunoglobulin homology <IMM>

Query Match 67.1%; Score 57; DB 2; Length 130;
Best Local Similarity 68.8%; Pred. No. 0.062;
Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVKG 17
||:||||:|||||
DB 58 MSSSGRTYYADSVKG 73

RESULT 10

I47191
Ig heavy chain variable VDJ region - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Jan-2000
C:Accession: I47191
R:Sun, J.; Kacsokovics, I.; Brown, W.R.; Butler, J.E.
J. Immunol. 153, 5618-5627, 1994
A:Title: Expressed swine VH genes belong to a small VH gene family homologous to human V
A:Reference number: I47177; MUID:95081609; PMID:7989761
A:Accession: I47191
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-132 <SUN>
A:Cross-references: EMBL:U15450; NID:G571386; PIDN:AAA67016.1; PID:G571387
C:Superfamily: immunoglobulin V region; immunoglobulin homology
F:34-117/Domain: immunoglobulin homology <IMM>

Query Match 67.1%; Score 57; DB 2; Length 132;
Best Local Similarity 68.8%; Pred. No. 0.063;
Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVKG 17
||:||||:|||||
DB 70 ISDSGRRTDYADSVKG 85

RESULT 11

PH0875
Ig heavy chain V region (anti-DNA, H2F) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 16-Aug-1996
C:Accession: PH0875
R:Manheimer-Loiry, A.; Katz, J.B.; Pillinger, M.; Ghossein, C.; Smith, A.; Diamond, B.
J. Exp. Med. 174, 1639-1652, 1991
A:Title: Molecular characteristics of antibodies bearing an anti-DNA-associated idiotype
A:Reference number: PH0862; MUID:92078875; PMID:1660528
A:Accession: PH0875
A:Molecule type: DNA
A:Residues: 1-97 <MAN>
C:Comment: This antibody is produced by Epstein-Barr virus-transformed B cell that bears
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:14-97/Domain: immunoglobulin homology <IMM>
F:30-35/Region: complementarity-determining 1
F:49-66/Region: complementarity-determining 2

Query Match 65.9%; Score 56; DB 2; Length 97;
Best Local Similarity 75.0%; Pred. No. 0.066;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVKG 17
||:||||:|||||
DB 50 ISSRGSTIYYADSVKG 65

RESULT 12

S36273
Ig heavy chain V region (clone alpha-THY-32) - human (fragment)
C:Species: Homo sapiens (man)

C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 23-Jul-1999
C:Accession: S36273
R:Griffiths, A.D.; Malmqvist, M.; Marks, J.D.; Bye, J.M.; Embleton, M.J.; McCafferty, J.;
EMBO J. 12, 725-734, 1993
A:Title: Human anti-self antibodies with high specificity from phage display libraries.
A:Reference number: S36256; MUID:93178448; PMID:7679990
A:Accession: S36273
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-120 <GRI>
A:Cross-references: EMBL:Z18834; NID:G33117; PIDN:CAA79286.1; PID:G939896
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 62.4%; Score 53; DB 2; Length 120;
Best Local Similarity 68.8%; Pred. No. 0.25;
Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVKG 17
||:||||:|||||
DB 51 ISSSGSIGYADSVKG 66

RESULT 13

S24251
Ig heavy chain V region (N54P3-D-JH4) - human
C:Species: Homo sapiens (man)
C:Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999
C:Accession: S24251
R:Stewart, A.K.; Huang, C.; Stollar, B.D.; Schwartz, R.S.
submitted to the EMBL Data Library, June 1992
A:Description: A single VH gene predominates in the rearranged and expressed human B cell
A:Reference number: S24247
A:Accession: S24251
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-108 <STE>
A:Cross-references: EMBL:X67072; NID:G38401; PIDN:CAA47457.1; PID:G38402
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:11-94/Domain: immunoglobulin homology <IMM>

Query Match 61.2%; Score 52; DB 2; Length 108;
Best Local Similarity 68.8%; Pred. No. 0.33;
Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVKG 17
||:||||:|||||
DB 47 ISSSGSAIYYADSVKG 62

RESULT 14

S36267
Ig heavy chain V region (clone alpha-THY-33) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 23-Jul-1999
C:Accession: S36267
R:Griffiths, A.D.; Malmqvist, M.; Marks, J.D.; Bye, J.M.; Embleton, M.J.; McCafferty, J.;
EMBO J. 12, 725-734, 1993
A:Title: Human anti-self antibodies with high specificity from phage display libraries.
A:Reference number: S36256; MUID:93178448; PMID:7679990
A:Accession: S36267
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-115 <GRI>
A:Cross-references: EMBL:Z18836; NID:G33117; PIDN:CAA79288.1; PID:G939897
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 61.2%; Score 52; DB 2; Length 115;
Best Local Similarity 62.5%; Pred. No. 0.35;

Matches 10; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVKG 17
||:|||||:
Db 51 ISYDGRSVYADSVQG 66

RESULT 15

IG heavy chain V region - human
A:Species: Homo sapiens (man)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 16-Aug-1996
C:Accession: S30532
R:Marlette, X.

submitted to the EMBL Data Library, October 1992
A:Reference number: S30520
A:Accession: S30532
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-123 <NAR>
C:Cross-references: EMBL:218318
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 61.2%; Score 52; DB 2; Length 123;
Best Local Similarity 68.8%; Pred. No. 0.38;
Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVKG 17
||:|||||:
Db 51 ISWNSGTLGYADSVKG 66

RESULT 16

IG heavy chain V region (DP-51) - human (fragment)
A:Species: Homo sapiens (man)
C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999
C:Accession: S26894
R:Tomlinson, I.M.; Walter, G.; Marks, J.D.; Llewellyn, M.B.; Winter, G.
J. Mol. Biol. 227, 776-798, 1992
A:Title: The repertoire of human germline V(H) sequences reveals about fifty groups of V
A:Reference number: S26885; MUID:93021117; PMID:1404388
A:Accession: S26894
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-98 <TOM>
C:Cross-references: EMBL:212351; NID:G32924; PIDN:CAA78221.1; PID:G32925
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 60.0%; Score 51; DB 2; Length 98;
Best Local Similarity 68.8%; Pred. No. 0.43;
Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVKG 17
||:|||||:
Db 51 ISSSSSTIYADSVKG 66

RESULT 17

IG heavy chain V region (clone alpha-FOG1-H6) - human (fragment)
A:Species: Homo sapiens (man)
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 23-Jul-1999
C:Accession: S36282
R:Griffiths, A.D.; Malmqvist, M.; Marks, J.D.; Bye, J.M.; Embleton, M.J.; McCafferty, J.
EMBO J. 12, 725-734, 1993
A:Title: Human anti-self antibodies with high specificity from phage display libraries.
A:Reference number: S36256; MUID:93178448; PMID:7679990
A:Accession: S36282

A>Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-110 <GRI>
C:Cross-references: EMBL:Z18824; NID:G33111; PIDN:CAA79276.1; PID:G939891
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 60.0%; Score 51; DB 2; Length 110;
Best Local Similarity 68.8%; Pred. No. 0.49;
Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVKG 17
||:|||||:
Db 51 ISSSSSTIYADSVKG 66

RESULT 18

M3HWE
IG heavy chain V-III region (WEA) - human (tentative sequence)
A:Species: Homo sapiens (man)
C:Date: 03-Aug-1984 #sequence_revision 03-Aug-1984 #text_change 31-Mar-2000
C:Accession: A02046
R:Goni, F.; Frangione, B.
Proc. Natl. Acad. Sci. U.S.A. 80, 4837-4841, 1983
A:Title: Amino acid sequence of the Fv region of a human monoclonal IgM (protein WEA) wit
A:Reference number: A91964; MUID:83273707; PMID:6410398
A:Accession: A02046
A:Molecule type: protein
A:Residues: 1-114 <GON>
C:Comment: This chain was obtained from a monoclonal antibody against 3,4-pyruvylated gal
C:Genetics:
A:Gene: GDB:IGHV@
A:Cross-references: GDB:128528; OMIM:147070
A:Map position: 14q32.33-14q32.33
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin; pyroglutamic acid
F:15-98/Domain: immunoglobulin homology <IMM>
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:22-96/Disulfide bonds: #status predicted

Query Match 60.0%; Score 51; DB 1; Length 114;
Best Local Similarity 68.8%; Pred. No. 0.5;
Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVKG 17
||:|||||:
Db 51 IGGSGSTIYADSVKG 66

RESULT 19

S21980
IG heavy chain V-gene (clone WHG26) precursor - human (fragment)
A:Species: Homo sapiens (man)
C:Date: 20-Feb-1995 #sequence_revision 25-Oct-1996 #text_change 23-Jul-1999
C:Accession: S60295; S21980
R:Kueppers, R.; Fischer, U.; Rajewsky, K.; Gause, A.
Immunol. Lett. 34, 57-62, 1992
A:Title: Immunoglobulin heavy and light chain gene sequences of a human CD5 positive imm
A:Reference number: S60295; MUID:93122853; PMID:1282498
A:Accession: S60295
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-117 <KUZ>
A:Cross-references: EMBL:X62130; NID:G38344; PIDN:CAA44061.1; PID:G38345
C:Genetics:
A:Introns: 16/1
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:34-117/Domain: immunoglobulin homology <IMM>

Query Match 60.0%; Score 51; DB 2; Length 117;
Best Local Similarity 68.8%; Pred. No. 0.52;

Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVKVG 17
||: || |||||
Db 70 ISSSSSTIYADSVKVG 85

RESULT 20

PH1661
IG heavy chain V region (clone SJI) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 16-Aug-1996
C:Accession: PH1661
R:Hallison, J.L.; Karr, N.S.; Opplinger, I.R.; Mannik, M.; Sasso, E.H.
J. Exp. Med. 178, 331-336, 1993
A:Title: The structural basis of germline-encoded VH3 immunoglobulin binding to staphylo
A:Reference number: PH1642; MUID:93301610; PMID:8315388
A:Accession: PH1661

A:Molecule type: mRNA

A:Residues: 1-121 <HIL>

A:Experimental source: B cell

C:Superfamily: immunoglobulin V region; immunoglobulin homology.

C:Keywords: heterotetramer; immunoglobulin

F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 60.0%; Score 51; DB 2; Length 121;
Best Local Similarity 64.7%; Pred. No. 0.54;
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 VIENGRNTINYADSVKVG 17
||: || |||||
Db 42 VISDDGSNKYADSVKVG 58

RESULT 21

S26794
IG heavy chain V region - human
C:Species: Homo sapiens (man)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 17-Mar-1999
C:Accession: S26794
R:Mortari, F.; Newton, J.A.; Wang, J.Y.; Schroeder Jr., H.W.
Eur. J. Immunol. 22, 241-245, 1992
A:Title: The human cord blood antibody repertoire. Frequent usage of the V(H)7 gene fami
A:Reference number: S26786; MUID:92111632; PMID:1730251
A:Accession: S26794

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-123 <MOR>

C:Cross-references: EMBL:X61011

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 60.0%; Score 51; DB 2; Length 123;
Best Local Similarity 68.8%; Pred. No. 0.54;
Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVKVG 17
||: || |||||
Db 51 ISSSSSTIYADSVKVG 66

RESULT 22

S30531
IG heavy chain V region - human
C:Species: Homo sapiens (man)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 16-Aug-1996
C:Accession: S30531
R:Marlette, X.

submitted to the EMBL Data Library, October 1992

A:Reference number: S30520

A:Accession: S30531

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-125 <MAP>

C:Cross-references: EMBL:Z18317

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 60.0%; Score 51; DB 2; Length 125;
Best Local Similarity 68.8%; Pred. No. 0.55;
Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVKVG 17
||: || |||||
Db 51 ISSSSSTIYADSVKVG 66

RESULT 23

S23624
IG heavy chain V region - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 23-Jul-1999
C:Accession: S23624
R:Oles, T.; Lu, E.W.; Huang, D.F.; Soto-Gil, R.W.; Defetos, M.; Kozin, F.; Carson, D.A.;
J. Exp. Med. 175, 831-842, 1992
A:Title: Genetic analysis of self-associating immunoglobulin G rheumatoid factors from tv

A:Reference number: S23623; MUID:92156804; PMID:1740665

A:Accession: S23624

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-143 <OLE>

C:Cross-references: EMBL:X59703; NID:G32012; PIDN:CAA42224.1; PID:G32013

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 60.0%; Score 51; DB 2; Length 143;
Best Local Similarity 68.8%; Pred. No. 0.63;
Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVKVG 17
||: || |||||
Db 51 ISSSSSTIYADSVKVG 66

RESULT 24

PH0874
IG heavy chain V region (anti-DNA, 1X7RG1) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 16-Aug-1996
C:Accession: PH0874
R:Manheimer-Lory, A.; Katz, J.B.; Pillinger, M.; Ghossein, C.; Smith, A.; Diamond, B.
J. Exp. Med. 174, 1639-1652, 1991
A:Title: Molecular characteristics of antibodies bearing an anti-DNA-associated idiotype.
A:Reference number: PH0862; MUID:92078875; PMID:1660528
A:Accession: PH0874

A:Molecule type: DNA

A:Residues: 1-98 <MAN>

C:Comment: This antibody is produced by Epstein-Barr virus-transformed B cell that bears

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:15-98/Domain: immunoglobulin homology <IMM>

F:31-36/Region: complementarity-determining 1

F:50-67/Region: complementarity-determining 2

Query Match 58.8%; Score 50; DB 2; Length 98;
Best Local Similarity 66.7%; Pred. No. 0.63;
Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 SENGRTINYADSVKVG 17
||: || |||||
Db 52 SSSGSTMYYADSVKVG 66

RESULT 25
S26927
Ig heavy chain V region (DP-31) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999
C:Accession: S26927
R:Tomlinson, I.M.; Walter, G.; Marks, J.D.; Llewelyn, M.B.; Winter, G.
J. Mol. Biol. 227, 776-798, 1992
A:Title: The repertoire of human germline V(H) sequences reveals about fifty groups of V
A:Reference number: S26885; MUID:93021117; PMID:1404388
A:Accession: S26927
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-98 <TOM>
A:Cross-references: EMBL:Z12333; NID:g32885; PIDN:CAA78203.1; PID:g32886
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 58.8%; Score 50; DB 2; Length 98;
Best Local Similarity 68.8%; Pred. No. 0.63;
Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 2 ISENGRTINYADSVKG 17
|||:|||||
Db 51 ISWNSGSGYADSVKG 66

Search completed: September 24, 2004, 01:54:34
Job time : 26.3396 secs

This Page Blank (uspto)

Result No.	Query			ID	Description
	Score	Match	Length		
1	85	100.0	249	4	AAB68087
2	67	78.8	120	2	AAR47042
3	65	76.5	17	5	AAU82607
4	63	74.1	17	6	ADA91299
5	63	74.1	17	6	ADA90004
6	63	74.1	17	6	ADA91201
7	63	74.1	17	6	ADA90434
8	63	74.1	17	6	ADA91199
9	63	74.1	17	6	ADA91301
10	63	74.1	115	7	ADL91350
11	63	74.1	115	7	ADL91351
12	63	74.1	115	8	ADP66612
13	63	74.1	115	8	ADP66613
14	63	74.1	115	8	ADP66606
15	63	74.1	115	8	ADP66609
16	63	74.1	115	8	ADP66611
17	63	74.1	115	8	ADP66610
18	63	74.1	115	8	ADP66605
19	63	74.1	115	8	ADP96608
20	63	74.1	115	8	ADP96610
21	63	74.1	115	8	ADP96607
22	63	74.1	115	8	ADP96611
23	63	74.1	115	8	ADP96603
24	63	74.1	115	8	ADP96604
25	63	74.1	115	8	ADP96609

XX Brodin TN, Karlstroem PJ, Ohlsson LG, Tordsson MJ, Kearney PP;
 PI Nilson BHK;
 XX WPI; 2001-308619/32.
 DR N-PSDB; AAF84797.
 XX Novel antibody for diagnosis, treatment of human metastatic and malignant
 PT diseases, has binding structure for target structure displayed on cell
 PT surface of human gastrointestinal epithelial tumor and normal cells.
 XX Claim 1; Page 55-56; 75pp; English.
 PS The present sequence represents a Monkey antibody light chain linked to a
 CC heavy chain. The antibody binds to a target structure displayed in and on
 CC the cell surface of human gastrointestinal epithelial tumour cells and in
 CC a subpopulation of normal human gastrointestinal epithelial cells. The
 CC target structure especially comprises alpha6beta4 integrin. This is a
 CC tumour-associated antigen. The antibody, and its fragments, are useful
 CC for treating conditions based on anti-angiogenic mechanism in humans.
 CC They are useful for treating human metastatic and malignant disease, for
 CC in vitro, in vivo diagnosis and prognosis of human malignant disease,
 CC comprising tumour typing, tumour screening, tumour diagnosis and
 CC prognosis and monitoring premalignant conditions. Quantitative in vivo
 CC diagnosis is carried out by determining the localization of antibody to
 CC tumour deposits in humans
 XX Sequence 249 AA;
 SQ Query Match 100.0%; Score 85; DB 4; Length 249;
 Best Local Similarity 100.0%; Pred. No. 2.7e-06; Mismatches 0; Indels 0; Gaps 0;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VISENGRTINYADSVKG 17
 Db 177 VISENGRTINYADSVKG 193
 |||||
 RESULT 2
 AAR47042
 ID AAR47042 standard; peptide; 120 AA.
 AC AAR47042;
 XX 25-MAR-2003 (revised)
 DT 02-SEP-1994 (first entry)
 XX Sequence of the consensus antibody variable domain of the heavy chain.
 DE Monoclonal antibody; heavy chain; variable domain; consensus.
 KW Synthetic.
 OS WO9404679-A1.
 PN 03-MAR-1994.
 XX 20-AUG-1993; 93WO-US007832.
 XX 21-AUG-1992; 92US-00934373.
 XX (GETH) GENENTECH INC.
 XX Carter PJ, Presta LG;
 XX WPI; 1994-083196/10.
 XX Preparation of improved humanised antibodies - by comparison of consensus
 PT and import complementarity determining regions and framework region
 PT sequences, e.g. to humanise murine.
 XX Claim 16; Page 108; 126pp; English.
 XX

CC The consensus variable domain sequences are derived from the most
 CC abundant subclasses in the sequence compilation of Kabat et al. Sequences
 CC of Proteins of Immunological Interest, National Institutes of Health,
 CC Bethesda MD (1987), namely, VL-kappa subgroup I and VH group III. In such
 CC IGG-gamma-1 human consensus sequences, the VI consensus domain has the AA
 CC sequence in AAR47041, and the VH consensus domain has the AA sequence in
 CC AAR47042. (Updated on 25-MAR-2003 to correct PN field.)
 XX Sequence 120 AA;
 SQ Query Match 78.8%; Score 67; DB 2; Length 120;
 Best Local Similarity 82.4%; Pred. No. 0.0015;
 Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 VISENGRTINYADSVKG 17
 Db 50 VISENGTTRYADSVKG 66
 |||||
 RESULT 3
 AAU82607
 ID AAU82607 standard; peptide; 17 AA.
 AC AAU82607;
 XX 29-AUG-2003 (revised)
 DT 23-APR-2002 (first entry)
 XX Llama CDR2 region variable heavy chain fragment #61.
 DE Llama phage display library; variable heavy domain fragment; VH; VH;
 KW sdab fragment; single domain anti-idiotypic antibody fragment;
 KW phage display technology; immune system response; CDR1/H1; CDR2; CDR3;
 KW complementarity determining region.
 XX Lama glama.
 OS WO200190190-A2.
 PN 29-NOV-2001.
 XX 25-MAY-2001; 2001WO-CA000763.
 XX 26-MAY-2000; 2000US-0207234P.
 XX (CANA) NAT RES COUNCIL CANADA.
 XX Tanha J, Dubuc G, Narang S;
 PI WPI; 2002-083093/11.
 DR New phage display library of variable heavy domain antigen-binding
 PT fragments derived from llama antibodies, useful for in vitro selection
 PT against any antigen of interest as a target.
 XX Disclosure; Page 19a; 46pp; English.
 XX The present invention relates to a phage display library of variable
 CC heavy domain (VH or VH) fragments (sdab fragments) derived from llama
 CC antibodies. The library is useful for in vitro selection against any
 CC antigen of interest as a target. Single domain anti-idiotypic antibody
 CC fragments are isolated from the library using phage display technology
 CC and an antibody serving as an antigen. Such anti-idiotypic antibody
 CC fragments have great potential in evoking the immune system response to
 CC pathological antigens and in vaccine development. The large size of the
 CC library considerably increases the probability of isolating from it
 CC antigen-binding fragments having high affinity to almost any
 CC predetermined target (antigen of interest). The library eliminates the
 CC development of anti-idiotypic antibodies by immunisation and allows
 CC isolation of anti-idiotypic antibodies. AAU82435-AAU82635 represent the
 CC llama heavy chain domain fragments of the invention. (Updated on 29-AUG-
 CC 2003 to standardise OS field)
 XX

SQ Sequence 17 AA;

Query Match 76.5%; Score 65; DB 5; Length 17;
Best Local Similarity 70.6%; Pred. No. 0.00033;
Matches 12; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 VISENGRTINYADSVKVG 17
::: ||| |||||

DB 1 LLRSGRITNYADSVKVG 17

RESULT 4

ADA91299 standard, peptide; 17 AA.

XX AC ADA91299;

XX DT 20-NOV-2003 (first entry)

XX DE MS-R Fab/antibody related peptide #347.

XX KW antibody molecule; antibody; beta-A4 peptide; Abeta4; neuroprotective;
KW neotropic; antiparkinsonian; gene therapy; amyloidogenesis;
KW amyloid-plaque formation; beta-amyloid plaque; immunisation; dementia;
KW Alzheimer's disease; motor neuropathy; Down's syndrome;
KW Creutzfeldt Jacob disease; hereditary cerebral haemorrhage; amyloidosis;
KW Parkinson's disease; HIV-related dementia; amyotrophic lateral sclerosis;
KW neuronal disorder; aging.

XX OS Synthetic.
OS Homo sapiens.

XX FN WO2003070760-A2.

XX PD 28-AUG-2003.

XX PF 20-FEB-2003; 2003WO-EP001759.

XX PR 20-FEB-2002; 2002EP-00003844.

XX PA (HOFF) HOFFMANN LA ROCHE & CO AG F.

XX PI Bardroff M, Bohrmann B, Brockhaus M, Huber W, Kretzschmar T;
PI Loehning C, Loetscher H, Nordstedt C, Rothe C;

XX DR WPI; 2003-663848/62.

XX PT New antibody molecule capable of specifically recognizing two regions of
PT the beta-A4 peptide, useful for diagnosing, preventing or treating
PT diseases associated with amyloidogenesis or amyloid-plaque formation
PT (e.g. dementia).

XX PS Disclosure; Page 189-190; 312pp; English.

XX CC The present invention describes an antibody molecule (I) capable of
CC specifically recognising two regions of the beta-A4 peptide/Abeta4. The
CC first region comprises the amino acid sequence Ala-Glu-Phe-Arg-His-Asp-

CC Ser-Gly-Tyr ADA89886 or its fragment, and the second region comprises the
 CC amino acid sequence Val-His-His-Gln-Lys-Leu-Val-Phe-Phe-Ala-Glu-Asp-Val-
 CC Gly ADA89887 or its fragment. Also described: (1) a nucleic acid molecule
 CC encoding (1); (2) a vector comprising the nucleic acid of (1); (3) a host
 CC cell comprising the vector of (2); (4) preparing (1), comprising
 CC culturing the host cell of (3) under conditions that allow synthesis of
 CC (1) and recovering (1) from the culture; (5) a composition comprising (1)
 CC or an antibody molecule produced by method (4); (6) a kit comprising (1),
 CC nucleic acid of (1), vector of (2) or host cell of (3); (7) optimising
 CC (1); (8) testing the resulting Fab optimisation library by panning
 CC against Abeta/Abeta4; (9) identifying optimised clones; (10) expressing
 CC of selected, optimised clones; (11) preparing a pharmaceutical
 CC composition, comprising optimisation of (1), and formulating the
 CC optimised antibody/antibody molecule with a carrier; and (12) a
 CC pharmaceutical composition prepared by method (8). (1) has
 CC neuroprotective, neurotropic and antiparkinsonian activities, and can be
 CC used in gene therapy. The antibody molecule (1), nucleic acid molecule,
 CC vector or host is useful in preparing a pharmaceutical composition for
 CC the prevention and/or treatment of a disease associated with
 CC amyloidogenesis and/or amyloid-plaque formation. The antibody molecule
 CC may also be used in preparing a diagnostic composition for the detection
 CC of the disease mentioned above. The antibody is used for the
 CC disintegration of beta-amyloid plaques or for passive immunisation
 CC against beta-amyloid plaque formation. In particular, the disease is
 CC dementia, Alzheimer's disease, motor neuropathy, Down's syndrome,
 CC Creutzfeldt Jacob disease, hereditary cerebral haemorrhage with
 CC amyloidosis Dutch type, Parkinson's disease, HIV-related dementia,
 CC amyotrophic lateral sclerosis or neuronal disorders related to aging. The
 CC present sequence is used in the exemplification of the present invention.

XX Sequence 17 AA;

Query Match 74.1%; Score 63; DB 6; Length 17;
 Best Local Similarity 76.5%; Pred. No. 0.00074;
 Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 VISENGRTINYADSVKG 17
 |||||:|:|||||||
 Db 1 VISETGKNIYADSVKG 17

RESULT 6

ID ADA91201 standard; peptide; 17 AA.

XX ADA91201;

XX 20-NOV-2003 (first entry)

DE MS-R Fab/antibody related peptide #249.

XX antibody molecule; antibody; beta-A4 peptide; Abeta4; neuroprotective;
 KW neurotropic; antiparkinsonian; gene therapy; amyloidogenesis;
 KW amyloid-plaque formation; beta-amyloid plaque; immunisation; dementia;
 KW Alzheimer's disease; motor neuropathy; Down's syndrome;
 KW Creutzfeldt Jacob disease; hereditary cerebral haemorrhage; amyloidosis;
 KW Parkinson's disease; HIV-related dementia; amyotrophic lateral sclerosis;
 KW neuronal disorder; aging.

OS Synthetic.

OS Homo sapiens.

XX WO2003070760-A2.

XX 28-AUG-2003.

XX 20-FEB-2003; 2003WO-BP001759.

XX 20-FEB-2002; 2002EP-00003844.

XX (HOFF) HOFFMANN LA ROCHE & CO AG F.
 PA (MORP-) MORPHOSYS AG.

PI

PI Bardroff M, Bohrmann B, Brockhaus M, Huber W, Kretzschmar T;

XX Loehning C, Loetscher H, Nordstedt C, Rothe C;

XX WPI; 2003-663848/62.

XX New antibody molecule capable of specifically recognizing two regions of
 PT the beta-A4 peptide, useful for diagnosing, preventing or treating
 PT diseases associated with amyloidogenesis or amyloid-plaque formation
 PT (e.g. dementia).

XX Example 13; Page 92; 312pp; English.

XX The present invention describes an antibody molecule (1) capable of
 CC specifically recognising two regions of the beta-A4 peptide/Abeta4. The
 CC first region comprises the amino acid sequence Ala-Glu-Phe-Arg-His-Asp-
 CC Ser-Gly-Tyr ADA89886 or its fragment, and the second region comprises the
 CC amino acid sequence Val-His-His-Gln-Lys-Leu-Val-Phe-Phe-Ala-Glu-Asp-Val-
 CC Gly ADA89887 or its fragment. Also described: (1) a nucleic acid molecule
 CC encoding (1); (2) a vector comprising the nucleic acid of (1); (3) a host
 CC cell comprising the vector of (2); (4) preparing (1), comprising
 CC culturing the host cell of (3) under conditions that allow synthesis of
 CC (1) and recovering (1) from the culture; (5) a composition comprising (1)
 CC or an antibody molecule produced by method (4); (6) a kit comprising (1),
 CC nucleic acid of (1), vector of (2) or host cell of (3); (7) optimising
 CC (1); (8) testing the resulting Fab optimisation library by panning
 CC against Abeta/Abeta4; (9) identifying optimised clones; (10) expressing
 CC of selected, optimised clones; (11) preparing a pharmaceutical
 CC composition, comprising optimisation of (1), and formulating the
 CC optimised antibody/antibody molecule with a carrier; and (12) a
 CC pharmaceutical composition prepared by method (8). (1) has
 CC neuroprotective, neurotropic and antiparkinsonian activities, and can be
 CC used in gene therapy. The antibody molecule (1), nucleic acid molecule,
 CC vector or host is useful in preparing a pharmaceutical composition for
 CC the prevention and/or treatment of a disease associated with
 CC amyloidogenesis and/or amyloid-plaque formation. The antibody molecule
 CC may also be used in preparing a diagnostic composition for the detection
 CC of the disease mentioned above. The antibody is used for the
 CC disintegration of beta-amyloid plaques or for passive immunisation
 CC against beta-amyloid plaque formation. In particular, the disease is
 CC dementia, Alzheimer's disease, motor neuropathy, Down's syndrome,
 CC Creutzfeldt Jacob disease, hereditary cerebral haemorrhage with
 CC amyloidosis Dutch type, Parkinson's disease, HIV-related dementia,
 CC amyotrophic lateral sclerosis or neuronal disorders related to aging. The
 CC present sequence is used in the exemplification of the present invention.

XX Sequence 17 AA;

Query Match 74.1%; Score 63; DB 6; Length 17;

Best Local Similarity 76.5%; Pred. No. 0.00074;

Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 VISENGRTINYADSVKG 17

|||||:|:|||||||
 Db 1 VISETGKNIYADSVKG 17

RESULT 7

ID ADA90434

AD A90434 standard; peptide; 17 AA.

XX ADA90434;

XX 20-NOV-2003 (first entry)

XX MS-Roche #3.4.H7 H-CDR2 amino acid sequence.

XX antibody molecule; antibody; beta-A4 peptide; Abeta4; neuroprotective;
 KW neurotropic; antiparkinsonian; gene therapy; amyloidogenesis;
 KW amyloid-plaque formation; beta-amyloid plaque; immunisation; dementia;
 KW Alzheimer's disease; motor neuropathy; Down's syndrome;
 KW Creutzfeldt Jacob disease; hereditary cerebral haemorrhage; amyloidosis;
 KW Parkinson's disease; HIV-related dementia; amyotrophic lateral sclerosis;
 KW neuronal disorder; aging.

XX Synthetic.
OS Homo sapiens.
PN WO2003070760-A2.
XX 28-AUG-2003.
XX 20-FEB-2003; 2003WO-EP001759.
XX 20-FEB-2002; 2002EP-00003844.
XX (HOFF) HOFFMANN LA ROCHE & CO AG F.
PA (MORP-) MORPHOSYS AG.
XX Bardroff M, Bohrmann B, Brockhaus M, Huber W, Kretzschmar T;
PI Loehning C, Loetscher H, Nordstedt C, Rothe C;
XX WPI; 2003-663848/62.
XX New antibody molecule capable of specifically recognising two regions of
PT the beta-A4 peptide, useful for diagnosing, preventing or treating
PT diseases associated with amyloidogenesis or amyloid-plaque formation
PT (e.g. dementia).
XX
PS Example 5; Page 64; 312pp; English.
CC The present invention describes an antibody molecule (I) capable of
CC specifically recognising two regions of the beta-A4 peptide/Abeta4. The
CC first region comprises the amino acid sequence Ala-Glu-Phe-Arg-His-Asp-
CC Ser-Gly-Tyr ADA89886 or its fragment, and the second region comprises the
CC amino acid sequence Val-His-His-Gln-Lys-Leu-Val-Phe-Phe-Ala-Glu-Asp-Val-
CC Gly ADA89887 or its fragment. Also described: (1) a nucleic acid molecule
CC encoding (1); (2) a vector comprising the nucleic acid of (1); (3) a host
CC cell comprising the vector of (2); (4) preparing (1), comprising
CC culturing the host cell of (3) under conditions that allow synthesis of
CC (1) and recovering (1) from the culture; (5) a composition comprising (1)
CC or an antibody molecule produced by method (4); (6) a kit comprising (1)
CC or an antibody molecule produced by method (4); (6) a kit comprising (1)
CC or an antibody molecule produced by method (4); (6) a kit comprising (1),
CC nucleic acid of (1), vector of (2) or host cell of (3); (7) optimising
CC (1); (8) testing the resulting Fab optimisation library by panning
CC against Abeta4/Abeta4; (9) identifying optimised clones; (10) expressing
CC of selected, optimised clones; (11) preparing a pharmaceutical
CC composition, comprising optimisation of (1), and formulating the
CC optimised antibody/molecule with a carrier; and (12) a
CC pharmaceutical composition prepared by method (8). (1) has
CC neuroprotective, neurotropic and antiparkinsonian activities, and can be
CC used in gene therapy. The antibody molecule (I), nucleic acid molecule,
CC vector or host is useful in preparing a pharmaceutical composition for
CC the prevention and/or treatment of a disease associated with
CC amyloidogenesis and/or amyloid-plaque formation. The antibody molecule
CC may also be used in preparing a diagnostic composition for the detection
CC of the disease mentioned above. The antibody is used for the
CC disintegration of beta-amyloid plaques or for passive immunisation
CC against beta-amyloid plaque formation. In particular, the disease is
CC dementia, Alzheimer's disease, motor neuropathy, Down's syndrome,
CC Creutzfeldt Jacob disease, hereditary cerebral haemorrhage with
CC amyloidosis Dutch type, Parkinson's disease, HIV-related dementia,
CC amyotrophic lateral sclerosis or neuronal disorders related to aging. The
CC present sequence is used in the exemplification of the present invention.

XX Sequence 17 AA;
SQ
Query Match 74.1%; Score 63; DB 6; Length 17;
Best Local Similarity 76.5%; Pred. No. 0.00074;
Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 VISEGRTINYADSVKG 17
||||| : |||||
Db 1 VISEGKNIYADSVKG 17
||||| : |||||

RESULT 8
ADA91199

ID ADA91199 standard; peptide; 17 AA.
XX
AC ADA91199;
XX
DT 20-NOV-2003 (first entry)
XX
DE MS-R Fab/antibody related peptide #247.
XX
KW antibody molecule; antibody; beta-A4 peptide; Abeta4; neuroprotective;
KW neurotropic; antiparkinsonian; gene therapy; amyloidogenesis;
KW amyloid-plaque formation; beta-amyloid plaque; immunisation; dementia;
KW Alzheimer's disease; motor neuropathy; Down's syndrome;
KW Creutzfeldt Jacob disease; hereditary cerebral haemorrhage; amyloidosis;
KW Parkinson's disease; HIV-related dementia; amyotrophic lateral sclerosis;
KW neuronal disorder; aging.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO2003070760-A2.
XX
PD 28-AUG-2003.
XX
XX 20-FEB-2003; 2003WO-EP001759.
XX
XX 20-FEB-2002; 2002EP-00003844.
XX (HOFF) HOFFMANN LA ROCHE & CO AG F.
PA (MORP-) MORPHOSYS AG.
XX Bardroff M, Bohrmann B, Brockhaus M, Huber W, Kretzschmar T;
PI Loehning C, Loetscher H, Nordstedt C, Rothe C;
XX WPI; 2003-663848/62.
XX New antibody molecule capable of specifically recognising two regions of
PT the beta-A4 peptide, useful for diagnosing, preventing or treating
PT diseases associated with amyloidogenesis or amyloid-plaque formation
PT (e.g. dementia).
XX
PS Example 13; Page 92; 312pp; English.
CC The present invention describes an antibody molecule (I) capable of
CC specifically recognising two regions of the beta-A4 peptide/Abeta4. The
CC first region comprises the amino acid sequence Ala-Glu-Phe-Arg-His-Asp-
CC Ser-Gly-Tyr ADA89886 or its fragment, and the second region comprises the
CC amino acid sequence Val-His-His-Gln-Lys-Leu-Val-Phe-Phe-Ala-Glu-Asp-Val-
CC Gly ADA89887 or its fragment. Also described: (1) a nucleic acid molecule
CC encoding (1); (2) a vector comprising the nucleic acid of (1); (3) a host
CC cell comprising the vector of (2); (4) preparing (1), comprising
CC culturing the host cell of (3) under conditions that allow synthesis of
CC (1) and recovering (1) from the culture; (5) a composition comprising (1)
CC or an antibody molecule produced by method (4); (6) a kit comprising (1)
CC or an antibody molecule produced by method (4); (6) a kit comprising (1),
CC nucleic acid of (1), vector of (2) or host cell of (3); (7) optimising
CC (1); (8) testing the resulting Fab optimisation library by panning
CC against Abeta4/Abeta4; (9) identifying optimised clones; (10) expressing
CC of selected, optimised clones; (11) preparing a pharmaceutical
CC composition, comprising optimisation of (1), and formulating the
CC optimised antibody/molecule with a carrier; and (12) a
CC pharmaceutical composition prepared by method (8). (1) has
CC neuroprotective, neurotropic and antiparkinsonian activities, and can be
CC used in gene therapy. The antibody molecule (I), nucleic acid molecule,
CC vector or host is useful in preparing a pharmaceutical composition for
CC the prevention and/or treatment of a disease associated with
CC amyloidogenesis and/or amyloid-plaque formation. The antibody molecule
CC may also be used in preparing a diagnostic composition for the detection
CC of the disease mentioned above. The antibody is used for the
CC disintegration of beta-amyloid plaques or for passive immunisation
CC against beta-amyloid plaque formation. In particular, the disease is
CC dementia, Alzheimer's disease, motor neuropathy, Down's syndrome,
CC Creutzfeldt Jacob disease, hereditary cerebral haemorrhage with
CC amyloidosis Dutch type, Parkinson's disease, HIV-related dementia,
CC amyotrophic lateral sclerosis or neuronal disorders related to aging. The
CC present sequence is used in the exemplification of the present invention.

CC present sequence is used in the exemplification of the present invention.

XX Sequence 17 AA;

SQ

Query Match 74.1%; Score 63; DB 6; Length 17;
Best Local Similarity 76.5%; Pred. No. 0.00074;
Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VISENGRTINYADSVKG 17
|||||:|||||||
DB 1 VISETGKNIYYADSVKG 17

RESULT 9

ADA91101

ID ADA91301 standard; peptide; 17 AA.

XX

AC ADA91301;

XX

DT 20-NOV-2003 (first entry)

XX

DE MS-R Fab/antibody related peptide #349.

XX

KW antibody molecule; antibody; beta-A4 peptide; Abeta4; neuroprotective;
KW neurotropic; antiparkinsonian; gene therapy; amyloidogenesis;
KW amyloid-plaque formation; beta-amyloid plaque; immunisation; dementia;
KW Alzheimer's disease; motor neuropathy; Down's syndrome;
KW Creutzfeldt Jacob disease; hereditary cerebral haemorrhage; amyloidosis;
KW Parkinson's disease; HIV-related dementia; amyotrophic lateral sclerosis;
KW neuronal disorder; aging.

XX

OS Synthetic.

OS Homo sapiens.

XX

PN WO2003070760-A2.

XX

PD 28-AUG-2003.

XX

PF 20-FEB-2003; 2003WO-EP001759.

XX

PR 20-FEB-2002; 2002EP-00003844.

XX

PA (HOFF) HOFFMANN LA ROCHE & CO AG F.

PA (MORP-) MORPHOSYS AG.

XX

PI Bardroff M, Bohrmann B, Brockhaus M, Huber W, Kretzschmar T;
PI Loehning C, Loetscher H, Nordstedt C, Rothe C;

XX

DR WPI; 2003-663848/62.

XX

PT New antibody molecule capable of specifically recognizing two regions of
PT the beta-A4 peptide, useful for diagnosing, preventing or treating
PT diseases associated with amyloidogenesis or amyloid-plaque formation
PT (e.g. dementia).

XX

PS Example 13; Page 93; 312pp; English.

XX

CC The present invention describes an antibody molecule (I) capable of
CC specifically recognising two regions of the beta-A4 peptide/Abeta4. The
CC first region comprises the amino acid sequence Ala-Glu-Phe-Arg-His-Asp-
CC Ser-Gly-Tyr ADA89886 or its fragment, and the second region comprises the
CC amino acid sequence Val-His-His-Gln-Lys-Leu-Val-Phe-Phe-Ala-Glu-Asp-Val-
CC Gly ADA89887 or its fragment. Also described: (1) a nucleic acid molecule
CC encoding (I); (2) a vector comprising the nucleic acid of (1); (3) a host
CC cell comprising the vector of (2); (4) preparing (I), comprising
CC culturing the host cell of (3) under conditions that allow synthesis of
CC (I) and recovering (I) from the culture; (5) a composition comprising (I)
CC or an antibody molecule produced by method (4); (6) a kit comprising (I),
CC nucleic acid of (1), vector of (2) or host cell of (3); (7) optimising
CC (1); (8) testing the resulting Fab optimisation library by panning
CC against Abeta4/Abeta4; (9) identifying optimised clones; (10) expressing
CC of selected, optimised clones; (11) preparing a pharmaceutical
CC composition, comprising optimisation of (I), and formulating the

CC optimised antibody/antibody molecule with a carrier; and (12) a
CC pharmaceutical composition prepared by method (8). (I) has
CC neuroprotective, neurotropic and antiparkinsonian activities, and can be
CC used in gene therapy. The antibody molecule (I), nucleic acid molecule,
CC vector or host is useful in preparing a pharmaceutical composition for
CC the prevention and/or treatment of a disease associated with
CC amyloidogenesis and/or amyloid-plaque formation. The antibody molecule
CC may also be used in preparing a diagnostic composition for the detection
CC of the disease mentioned above. The antibody is used for the
CC disintegration of beta-amyloid plaques or for passive immunisation
CC against beta-amyloid plaque formation. In particular, the disease is
CC dementia, Alzheimer's disease, motor neuropathy, Down's syndrome,
CC Creutzfeldt Jacob disease, hereditary cerebral haemorrhage with
CC amyloidosis Dutch type, Parkinson's disease, HIV-related dementia,
CC amyotrophic lateral sclerosis or neuronal disorders related to aging. The
CC present sequence is used in the exemplification of the present invention.

XX

SQ Sequence 17 AA;

Query Match 74.1%; Score 63; DB 6; Length 17;
Best Local Similarity 76.5%; Pred. No. 0.00074;
Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VISENGRTINYADSVKG 17
|||||:|||||||
DB 1 VISETGKNIYYADSVKG 17

RESULT 10

ADL91350

ID ADL91350 standard; protein; 115 AA.

XX

AC ADL91350;

XX

DT 20-MAY-2004 (first entry)

XX

DE VH chain clone J48 of an intracellularly binding immunoglobulin SeqID 41.

XX

KW antibody; variable chain; cytostatic; cytoplasmic degradation;
KW intracellular relocation; specific antigen positive cancer; leukaemia;
KW lymphoma; intracellularly binding immunoglobulin; ras.

XX

OS Unidentified.

XX

PN WO2003077945-A1.

XX

PD 25-SEP-2003.

XX

PF 14-MAR-2003; 2003WO-GB001077.

XX

PR 14-MAR-2002; 2002GB-00006043.

PR 15-NOV-2002; 2002GB-00026723.

PR 15-NOV-2002; 2002GB-00026727.

XX

PA (MEDI-) MEDICAL RES COUNCIL.

XX

PI Lobato-Caballero MN, Rabbitts TH;

XX

DR WPI; 2003-779088/73.

XX

PT Use of an intracellularly binding immunoglobulin comprising at least one
PT antibody variable chain, in preparing a medicament for degrading one or
PT more specific antigens, or for treating specific antigen positive cancer,
PT e.g. leukemia.

XX

PS Example 7; SEQ ID NO 41; 86pp; English.

XX

CC This invention relates to novel immunoglobulin molecules that comprise at
CC least one antibody variable chain VH or VL framework region and are
CC capable of binding to a specific antigens within an intracellular
CC environment. Specifically, it refers to antibodies that can form an
CC insoluble complex with a cognate antigen, such that it can then be target
CC for degradation via the lysosome or proteosome systems. The present

CC invention describes the specific target antigen as the oncogenic fusion
CC protein BCR-ABL or the RAS antigen, such that this method can be used to
CC prepare a cytostatic medicament for the cytoplasmic degradation or
CC intracellular relocation of such an antigen or for the treatment of the
CC specific antigen positive cancer i.e. leukaemia or lymphoma. Furthermore,
CC the immunoglobulins may also be used for therapeutic, prophylactic or
CC diagnostic applications both in vitro and in vivo, as well as for assay
CC and reagent applications or in functional genomics. This polypeptide
CC sequence is a variable heavy chain (VH) framework region of an
CC intracellularly binding anti-ras antibody of the invention.

XX Sequence 115 AA;

Query Match 74.1%; Score 63; DB 7; Length 115;
Best Local Similarity 81.2%; Pred. No. 0.0073;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVKG 17
DB 51 ISSSGRTIYADSVKG 66
||:|||||

RESULT 11

ADL91351
ID ADL91351 standard; protein; 115 AA.

XX AC ADL91351;

XX DT 20-MAY-2004 (first entry)

XX DE VH chain clone 33 of an intracellularly binding immunoglobulin SeqID 42.

XX antibody; variable chain; cytostatic; cytoplasmic degradation;

XX intracellular relocation; specific antigen positive cancer; leukaemia;

XX lymphoma; intracellularly binding immunoglobulin; ras.

XX OS Unidentified.

XX PN WO2003077945-A1.

XX PD 25-SEP-2003.

XX PF 14-MAR-2003; 2003WO-GB001077.

XX PR 14-MAR-2002; 2002GB-00006043.

XX PR 15-NOV-2002; 2002GB-00026723.

XX PR 15-NOV-2002; 2002GB-00026727.

XX PA (MEDI-) MEDICAL RES COUNCIL.

XX PI Lobato-Caballero MN, Rabbitts TH;

XX DR WPI; 2003-779088/73.

XX Use of an intracellularly binding immunoglobulin comprising at least one
PT antibody variable chain, in preparing a medicament for degrading one or
PT more specific antigens, or for treating specific antigen positive cancer,
PT e.g. leukemia.

XX Example 7; SEQ ID NO 42; 86pp; English.

XX This invention relates to novel immunoglobulin molecules that comprise at
CC least one antibody variable chain VH or VL framework region and are
CC capable of binding to a specific antigens within an intracellular
CC environment. Specifically, it refers to antibodies that can form an
CC insoluble complex with a cognate antigen, such that it can then be target
CC for degradation via the lysosome or proteosome systems. The present
CC invention describes the specific target antigen as the oncogenic fusion
CC protein BCR-ABL or the RAS antigen, such that this method can be used to
CC prepare a cytostatic medicament for the cytoplasmic degradation or
CC intracellular relocation of such an antigen or for the treatment of the
CC specific antigen positive cancer i.e. leukaemia or lymphoma. Furthermore,
CC the immunoglobulins may also be used for therapeutic, prophylactic or

CC diagnostic applications both in vitro and in vivo, as well as for assay
CC and reagent applications or in functional genomics. This polypeptide
CC sequence is a variable heavy chain (VH) framework region of an
CC intracellularly binding anti-ras antibody of the invention.

XX Sequence 115 AA;

Query Match 74.1%; Score 63; DB 7; Length 115;
Best Local Similarity 81.2%; Pred. No. 0.0073;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVKG 17
DB 51 ISSSGRTIYADSVKG 66
||:|||||

RESULT 12

ADP66612
ID ADP66612 standard; protein; 115 AA.

XX AC ADP66612;

XX DT 26-AUG-2004 (first entry)

XX DE Anti-RAS intracellular mutated scFV Con33 VH chain, SEQ ID 9.

XX RAS; immunoglobulin; anti-RAS antibody; scFV.

XX OS Unidentified.

XX PN WO2004046187-A2.

XX PD 03-JUN-2004.

XX PF 14-NOV-2003; 2003WO-GB004944.

XX PR 15-NOV-2002; 2002GB-00026723.

XX PA (MEDI-) MEDICAL RES COUNCIL.

XX PI Rabbitts TH, Tanaka T;

XX DR WPI; 2004-431948/40.

XX New immunoglobulin molecule comprising at least one VH framework region
PT amino acid sequence, useful for the specific binding of a ligand within
PT an in vitro environment.

XX Example 3; Fig 3; 67pp; English.

XX The invention relates to an immunoglobulin molecule comprising at least
CC one VH framework region amino acid sequence showing at least 85 %
CC identity with the framework region consensus sequence (I), and/or a
CC molecule having at least one VH framework region amino acid sequence
CC showing at least 85 % identity with (I), where one or more of amino acids
CC 22 and 92 are not cysteine residues. The immunoglobulin suitable for in
CC vitro use can be produced by a library, generated using any one or more
CC of the VH framework region amino acid sequences cited above. The antibody
CC molecule is of a single variable domain type antibody. The single
CC variable domain type antibody is a heavy or light chain variable domain
CC antibody. The antibody comprises at least light and at least heavy chain
CC variable domain. The immunoglobulin molecule is useful for the specific
CC binding of a ligand within an in vitro environment. Sequences ADP66604-
CC ADP66613 represent variable domain heavy chain (VH) sequences of anti-RAS
CC intracellular scFVs.

XX Sequence 115 AA;

Query Match 74.1%; Score 63; DB 8; Length 115;

Best Local Similarity 81.2%; Pred. No. 0.0073;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVKG 17

Db	51	ISSSGRTIYADSVKG 66	:
RESULT 13			
ADP66613			
ID	ADP66613	standard; protein; 115 AA.	
XX	AC		
XX	ADP66613;		
XX	26-AUG-2004	(first entry)	
DT			
DE	Anti-RAS mutated scFV I21R33 (VHC22S:C92S) VH chain, SEQ ID 10.		
XX	RAS; immunoglobulin; anti-RAS antibody; scFV.		
XX	Unidentified.		
OS			
XX	WO2004046187-A2.		
XX	03-JUN-2004.		
XX	14-NOV-2003; 2003WO-GB004944.		
XX	15-NOV-2002; 2002GB-00026723.		
PR			
XX	(MEDI-) MEDICAL RES COUNCIL.		
FA			
XX	Rabbits TH, Tanaka T;		
XX	WPI; 2004-431948/40.		
XX	New immunoglobulin molecule comprising at least one VH framework region amino acid sequence, useful for the specific binding of a ligand within an in vitro environment.		
PT			
XX	Example 3; Fig 3; 67pp; English.		
XX	The invention relates to an immunoglobulin molecule comprising at least one VH framework region amino acid sequence showing at least 85 % identity with the framework region consensus sequence (I), and/or a molecule having at least one VH framework region amino acid sequence showing at least 85 % identity with (I), where one or more of amino acids 22 and 92 are not cysteine residues. The immunoglobulin suitable for in vitro use can be produced by a library, generated using any one or more of the VH framework region amino acid sequences cited above. The antibody molecule is of a single variable domain type antibody. The single variable domain type antibody is a heavy or light chain variable domain antibody. The antibody comprises at least light and at least heavy chain binding of a ligand within an in vitro environment. Sequences ADP66604-ADP66613 represent variable domain heavy chain (VH) sequences of anti-RAS intracellular scFVs.		
XX	Query Match	74.1%; Score 63; DB 8; Length 115;	
XX	Best Local Similarity	81.2%; Pred. No. 0.0073;	
XX	Matches	13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;	
QY	2	ISENGRTIYADSVKG 17	
DB	51	ISSSGRTIYADSVKG 66	
XX	RESULT 15		
ADP66609			
ID	ADP66609	standard; protein; 115 AA.	
XX	AC		
XX	ADP66609;		
XX	26-AUG-2004	(first entry)	
DT			
DE	Anti-RAS intracellular mutated scFV I21K33 VH chain, SEQ ID 6.		
XX	RAS; immunoglobulin; anti-RAS antibody; scFV.		
XX	Unidentified.		
OS			
XX	WO2004046187-A2.		
XX	03-JUN-2004.		
XX	14-NOV-2003; 2003WO-GB004944.		
XX	15-NOV-2002; 2002GB-00026723.		
PR			
XX	(MEDI-) MEDICAL RES COUNCIL.		
FA			
XX	Rabbits TH, Tanaka T;		
XX	WPI; 2004-431948/40.		
XX	New immunoglobulin molecule comprising at least one VH framework region amino acid sequence, useful for the specific binding of a ligand within an in vitro environment.		
PT			
XX	Example 3; Fig 3; 67pp; English.		
XX	The invention relates to an immunoglobulin molecule comprising at least one VH framework region amino acid sequence showing at least 85 % identity with the framework region consensus sequence (I), and/or a molecule having at least one VH framework region amino acid sequence showing at least 85 % identity with (I), where one or more of amino acids 22 and 92 are not cysteine residues. The immunoglobulin suitable for in vitro use can be produced by a library, generated using any one or more of the VH framework region amino acid sequences cited above. The antibody molecule is of a single variable domain type antibody. The single variable domain type antibody is a heavy or light chain variable domain antibody. The antibody comprises at least light and at least heavy chain binding of a ligand within an in vitro environment. Sequences ADP66604-ADP66613 represent variable domain heavy chain (VH) sequences of anti-RAS intracellular scFVs.		
XX	Query Match	74.1%; Score 63; DB 8; Length 115;	
XX	Best Local Similarity	81.2%; Pred. No. 0.0073;	
XX	Matches	13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;	
QY	2	ISENGRTIYADSVKG 17	
DB	51	ISSSGRTIYADSVKG 66	
XX	RESULT 14		
ADP66606			
ID	ADP66606	standard; protein; 115 AA.	
XX	AC		
XX	ADP66606;		
XX	26-AUG-2004	(first entry)	
DT			
DE	Anti-RAS intracellular mutated scFV I21K33 VH chain, SEQ ID 6.		
XX	RAS; immunoglobulin; anti-RAS antibody; scFV.		
XX	Unidentified.		
OS			
XX	WO2004046187-A2.		
XX	03-JUN-2004.		
XX	14-NOV-2003; 2003WO-GB004944.		
XX	15-NOV-2002; 2002GB-00026723.		
PR			
XX	(MEDI-) MEDICAL RES COUNCIL.		
FA			
XX	Rabbits TH, Tanaka T;		
XX	WPI; 2004-431948/40.		
XX	New immunoglobulin molecule comprising at least one VH framework region amino acid sequence, useful for the specific binding of a ligand within an in vitro environment.		
PT			
XX	Example 3; Fig 3; 67pp; English.		
XX	The invention relates to an immunoglobulin molecule comprising at least one VH framework region amino acid sequence showing at least 85 % identity with the framework region consensus sequence (I), and/or a molecule having at least one VH framework region amino acid sequence showing at least 85 % identity with (I), where one or more of amino acids 22 and 92 are not cysteine residues. The immunoglobulin suitable for in vitro use can be produced by a library, generated using any one or more of the VH framework region amino acid sequences cited above. The antibody molecule is of a single variable domain type antibody. The single variable domain type antibody is a heavy or light chain variable domain antibody. The antibody comprises at least light and at least heavy chain binding of a ligand within an in vitro environment. Sequences ADP66604-ADP66613 represent variable domain heavy chain (VH) sequences of anti-RAS intracellular scFVs.		
XX	Query Match	74.1%; Score 63; DB 8; Length 115;	
XX	Best Local Similarity	81.2%; Pred. No. 0.0073;	</

```

PR 15-NOV-2002; 2002GB-00026723.
XX (MEDI-) MEDICAL RES COUNCIL.
XX
XX PI Rabbitts TH, Tanaka T;
XX WPI; 2004-431948/40.
XX
XX New immunoglobulin molecule comprising at least one VH framework region
PT amino acid sequence, useful for the specific binding of a ligand within
PT an in vitro environment.
XX
XX Example 3; Fig 3; 67pp; English.
XX
XX The invention relates to an immunoglobulin molecule comprising at least
CC one VH framework region amino acid sequence showing at least 85 %
CC identity with the framework region consensus sequence (I), and/or a
CC molecule having at least one VH framework region amino acid sequence
CC showing at least 85 % identity with (I), where one or more of amino acids
CC 22 and 92 are not cysteine residues. The immunoglobulin suitable for in
CC vitro use can be produced by a library, generated using any one or more
CC of the VH framework region amino acid sequences cited above. The antibody
CC molecule is of a single variable domain type antibody. The single
CC variable domain type antibody is a heavy or light chain variable domain
CC antibody. The antibody comprises at least light and at least heavy chain
CC variable domain. The immunoglobulin molecule is useful for the specific
CC binding of a ligand within an in vitro environment. Sequences ADP66604-
CC ADP66613 represent variable domain heavy chain (VH) sequences of anti-RAS
CC intracellular scFVs.
XX
SQ Sequence 115 AA;
    Query Match          74.1%; Score 63; DB 8; Length 115;
    Best Local Similarity 81.2%; Pred. No. 0.0073;
    Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVKVG 17
   |||:|||||
Db 51 ISSGRTIYYADSVKVG 66

```

Search completed: April 28, 2005, 18:17:28
 Job time : 129.031 secs

This Page Blank (uap10)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2005, 18:06:50 ; Search time 34.5312 Seconds
(without alignments)
36.750 Million cell updates/sec

Title: US-10-088-639A-2_COPY_177_193
Perfect score: 85
Sequence: 1 VISENGRTINADVSKG 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	60	70.6	120	2	US-07-934-373C-4
2	60	70.6	120	3	US-08-437-642B-4
3	60	70.6	120	4	US-08-146-206C-4
4	60	70.6	120	4	US-09-705-686-4
5	60	70.6	120	4	US-09-705-392A-4
6	60	70.6	120	4	US-09-705-398-4
7	60	70.6	120	5	PCT-US93-07832-4
8	59	69.4	129	3	US-09-240-274-143
9	57	67.1	22	3	US-08-918-148-16
10	57	67.1	22	4	US-09-138-091A-16
11	57	67.1	98	2	US-08-665-202-31
12	57	67.1	98	4	US-09-315-574-31
13	57	67.1	117	3	US-08-545-809A-99
14	57	67.1	123	2	US-08-665-202-30
15	57	67.1	123	4	US-09-315-574-30
16	57	67.1	245	3	US-08-918-148-75
17	57	67.1	245	4	US-09-138-091A-73
18	56	65.9	124	3	US-09-240-274-4
19	54	63.5	17	4	US-09-192-854-128
20	54	63.5	124	3	US-09-240-274-2
21	54	63.5	124	3	US-09-240-274-3
22	54	63.5	124	3	US-09-240-274-7
23	53	62.4	116	3	US-08-545-809A-143
24	52	61.2	17	1	US-08-264-093-22
25	52	61.2	118	2	US-08-652-816A-11
26	52	61.2	120	1	US-08-264-093-14
27	52	61.2	179	3	US-08-862-124-2

28	52	61.2	268	4	US-09-976-118-1	Sequence 1, Appli
29	52	61.2	287	3	US-08-862-124-17	Sequence 17, Appli
30	52	61.2	304	3	US-08-862-124-14	Sequence 14, Appli
31	51	60.0	17	4	US-09-192-854-99	Sequence 99, Appli
32	51	60.0	17	4	US-09-424-840B-72	Sequence 72, Appli
33	51	60.0	17	4	US-09-424-840B-88	Sequence 88, Appli
34	51	60.0	98	1	US-08-211-202-116	Sequence 116, App
35	51	60.0	110	1	US-08-211-202-117	Sequence 117, App
36	51	60.0	111	1	US-08-211-202-134	Sequence 134, App
37	51	60.0	114	2	US-08-652-558-52	Sequence 52, Appli
38	51	60.0	117	3	US-08-545-809A-130	Sequence 130, Appli
39	51	60.0	121	4	US-09-553-949-7	Sequence 7, Appli
40	51	60.0	121	4	US-09-530-139-60	Sequence 60, Appli
41	51	60.0	121	4	US-09-424-840B-26	Sequence 26, Appli
42	51	60.0	126	1	US-08-478-039-95	Sequence 95, Appli
43	51	60.0	126	1	US-08-476-349A-95	Sequence 95, Appli
44	51	60.0	126	3	US-08-983-607-48	Sequence 48, Appli
45	51	60.0	137	1	US-08-331-398A-61	Sequence 61, Appli

ALIGNMENTS

RESULT 1
US-07-934-373C-4
; Sequence 4, Application US/07934373C
; Patent No. 5821337
; GENERAL INFORMATION:
; APPLICANT: Paul J. Carter
; APPLICANT: Leonard G. Presta
; TITLE OF INVENTION: Immunoglobulin Variants
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/934,373C
; FILING DATE: 21-Aug-1992
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/05126
; FILING DATE: 15-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/715272
; FILING DATE: 14-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P0709P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1994
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 120 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-07-934-373C-4

Query Match 70.6%; Score 60; DB 2; Length 120;
Best Local Similarity 76.5%; Fred. No. 0.012;
Matches 13; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 VISENGRTINADVSKG 17

```
Db          ||||| |||||
50 VISENGSDTYADSVKG 66

RESULT 2
US-08-437-642B-4
; Sequence 4, Application US/08437642B
; Patent No. 6054297
; GENERAL INFORMATION:
; APPLICANT: Paul J. Carter
; APPLICANT: Leonard G. Presta
; TITLE OF INVENTION: Immunoglobulin Variants
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Winpatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/437,642B
; FILING DATE: 09-May-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/934373
; FILING DATE: 21-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/146206
; FILING DATE: 17-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/05126
; FILING DATE: 15-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/715272
; FILING DATE: 14-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P07092C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1994
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 120 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-08-437-642B-4

Query Match          70.6%; Score 60; DB 3; Length 120;
Best Local Similarity 76.5%; Pred. No. 0.012;
Matches 13; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY          1 VISENGRTINYADSVKG 17
          ||||| |||||
Db          50 VISENGSDTYADSVKG 66

RESULT 4
US-09-705-686-4
; Sequence 4, Application US/09705686
; Patent No. 6639055
; GENERAL INFORMATION:
; APPLICANT: Carter, Paul J.
; APPLICANT: Presta, Leonard G.
; TITLE OF INVENTION: Method for Making Humanized Antibodies
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Winpatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/705,686
; FILING DATE: 02-NO. 6639055-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/146206
; FILING DATE: 17-NOV-1993
; APPLICATION NUMBER: 07/715272
; FILING DATE: 14-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P07092C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1994
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 120 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-08-437-642B-4

Query Match          70.6%; Score 60; DB 3; Length 120;
Best Local Similarity 76.5%; Pred. No. 0.012;
Matches 13; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY          1 VISENGRTINYADSVKG 17
          ||||| |||||
Db          50 VISENGSDTYADSVKG 66

RESULT 3
US-08-146-206C-4
; Sequence 4, Application US/08146206C
; Patent No. 6407213
; GENERAL INFORMATION:
; APPLICANT: Carter, Paul J.
; APPLICANT: Presta, Leonard G.
; TITLE OF INVENTION: Method for Making Humanized Antibodies
; NUMBER OF SEQUENCES: 26
```

NAME: Lee, Wendy M.
REGISTRATION NUMBER: 40,378
REFERENCE/DOCKET NUMBER: P0709PID3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-1994
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 120 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-705-686-4

Query Match 70.6%; Score 60; DB 4; Length 120;
Best Local Similarity 76.5%; Pred. No. 0.012;
Matches 13; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 VISENGRTINYADSVKG 17
|||||
Db 50 VISENGSDTYADSVKG 66

RESULT 5
US-09-705-392A-4
; Sequence 4, Application US/09705392A
; Patent No. 6719971
; GENERAL INFORMATION:
; APPLICANT: Carter, Paul J.

; TITLE OF INVENTION: Method for Making Humanized Antibodies
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/705,392A
FILING DATE: 02-NO. 6719971-2002
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/146206
FILING DATE: 17-NOV-1993
APPLICATION NUMBER: 07/715272
FILING DATE: 14-JUN-1991

ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 40,378
REFERENCE/DOCKET NUMBER: P0709PID1 REVISED
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-1994
TELEFAX: 650/952-9881

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 120 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-705-392A-4

Query Match 70.6%; Score 60; DB 4; Length 120;
Best Local Similarity 76.5%; Pred. No. 0.012;
Matches 13; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 VISENGRTINYADSVKG 17

Db 50 VISENGSDTYADSVKG 66
|||||
RESULT 6
US-09-705-398-4
; Sequence 4, Application US/09705398
; Patent No. 6800738
; GENERAL INFORMATION:
; APPLICANT: Carter, Paul J.
; TITLE OF INVENTION: Method for Making Humanized Antibodies
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/705,398
FILING DATE: 02-NO. 6800738-2000
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/146206
FILING DATE: 17-NOV-1993
APPLICATION NUMBER: 07/715272
FILING DATE: 14-JUN-1991

ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 40,378
REFERENCE/DOCKET NUMBER: P0709PID2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-1994
TELEFAX: 650/952-9881

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 120 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-705-398-4

Query Match 70.6%; Score 60; DB 4; Length 120;
Best Local Similarity 76.5%; Pred. No. 0.012;
Matches 13; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 VISENGRTINYADSVKG 17
|||||
Db 50 VISENGSDTYADSVKG 66

RESULT 7
PCT-US93-07832-4
; Sequence 4, Application PC/TUS9307832
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; TITLE OF INVENTION: Immunoglobulin Variants
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:

```
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/07832
; FILING DATE: 19930820
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/715272
; FILING DATE: 14-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/05126
; FILING DATE: 15-JUN-1992
; APPLICATION DATA:
; APPLICATION NUMBER: 07/934373
; FILING DATE: 21-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME:
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER: 709P2PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE:
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 120 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; PCT-US93-07832-4

Query Match 70.6%; Score 60; DB 5; Length 120;
Best Local Similarity 76.5%; Pred. No. 0.012;
Matches 13; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 VISENGRTINYADSVKG 17
Db 50 VISENGGYTRYADSVKG 66

RESULT 8
US-09-240-274-143
; Sequence 143, Application US/09240274
; Patent No. 6255455
; GENERAL INFORMATION:
; APPLICANT: Siegel, Donald L.
; TITLE OF INVENTION: Rh(D)-BINDING PROTEINS AND MAGNETICALLY ACTIVATED CELL
; FILE REFERENCE: 09596-42U2
; CURRENT FILING DATE: 1999-01-29
; EARLIER APPLICATION NUMBER: 60/081,380
; EARLIER FILING DATE: 1998-04-10
; EARLIER APPLICATION NUMBER: 60/028,550
; EARLIER FILING DATE: 1996-10-11
; NUMBER OF SEQ ID NOS: 224
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 143
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: anti-Rh(D) antibody clone SH20
US-09-240-274-143

Query Match 69.4%; Score 59; DB 3; Length 129;
Best Local Similarity 76.5%; Pred. No. 0.019;
Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VISENGRTINYADSVKG 17
Db 51 VISYDGSITYYADSVKG 67

; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/07832
; FILING DATE: 19930820
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/715272
; FILING DATE: 14-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/05126
; FILING DATE: 15-JUN-1992
; APPLICATION DATA:
; APPLICATION NUMBER: 07/934373
; FILING DATE: 21-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME:
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER: 709P2PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE:
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 120 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; PCT-US93-07832-4

Query Match 67.1%; Score 57; DB 3; Length 22;
Best Local Similarity 75.0%; Pred. No. 0.0056;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVKG 17
Db 2 ISSSGSTIYADSVKG 17

RESULT 10
US-09-138-091A-16
; Sequence 16, Application US/09138091A
; Patent No. 6737249
; GENERAL INFORMATION:
; APPLICANT: Adams, Camellia W.
; APPLICANT: Carter, Paul J.
; APPLICANT: Fendly, Brian M.
; APPLICANT: Gurney, Austin L.
; TITLE OF INVENTION: Agonist Antibodies
; FILE REFERENCE: 9491-013-27
; CURRENT APPLICATION NUMBER: US/09/138,091A
; CURRENT FILING DATE: 1998-08-21
; PRIOR APPLICATION NUMBER: US 60/056,736
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-09-138-091A-16

Query Match 67.1%; Score 57; DB 4; Length 22;
Best Local Similarity 75.0%; Pred. No. 0.0056;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVKG 17
Db 2 ISSSGSTIYADSVKG 17

RESULT 11
US-08-665-202-31
; Sequence 31, Application US/08665202
; Patent No. 5977322
; GENERAL INFORMATION:
```

APPLICANT: Marks, James D.
APPLICANT: Schier, Robert
TITLE OF INVENTION: No. 5977322el High Affinity Human Antibodies to
TITLE OF INVENTION: Tumor Antigens
NUMBER OF SEQUENCES: 141
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US 08/665,202
FILING DATE: 13-JUN-1996
CLASSIFICATION: 424

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/000,238
FILING DATE: 14-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/000,250
FILING DATE: 15-JUN-1995

ATTORNEY/AGENT INFORMATION:
NAME: Hunter, Tom
REGISTRATION NUMBER: 38,498
REFERENCE/DOCKET NUMBER: 02307E-061410

TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 31:

SEQUENCE CHARACTERISTICS:
LENGTH: 98 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide

US-08-665-202-31

Query Match 67.1%; Score 57; DB 2; Length 98;
Best Local Similarity 75.0%; Pred. No. 0.031;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVKG 17
||:|||||
DB 51 ISSSGSTIYADSVKG 66

RESULT 12
US-09-315-574-31
Sequence 31, Application US/09315574
Patent No. 6512097

GENERAL INFORMATION:
APPLICANT: Marks, James D.
APPLICANT: Schier, Robert

TITLE OF INVENTION: No. 6512097el High Affinity Human Antibodies to
TITLE OF INVENTION: Tumor Antigens
NUMBER OF SEQUENCES: 141
CORRESPONDENCE ADDRESS:

ADDRESSEE: Majestic, Parsons, Siebert & Haue P.C.
STREET: Four Embarcadero Center, Suite 1100
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4106

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/315,574
FILING DATE: 20-MAY-99
CLASSIFICATION: 530

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/000,238
FILING DATE: 14-JUN-1995

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/000,250
FILING DATE: 15-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/665,202
FILING DATE: 13-JUN-1996

ATTORNEY/AGENT INFORMATION:
NAME: Hunter, Tom
REGISTRATION NUMBER: 38,498

REFERENCE/DOCKET NUMBER: 02307E-061411
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 31:

SEQUENCE CHARACTERISTICS:
LENGTH: 98 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide

US-09-315-574-31

Query Match 67.1%; Score 57; DB 4; Length 98;
Best Local Similarity 75.0%; Pred. No. 0.031;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVKG 17
||:|||||
DB 51 ISSSGSTIYADSVKG 66

RESULT 13

US-08-545-809A-99
Sequence 99, Application US/08545809A
Patent No. 6096878

GENERAL INFORMATION:
APPLICANT: Honjo, Tasuku
APPLICANT: Matsuda, Fumihiko

TITLE OF INVENTION: HUMAN IMMUNOGLOBULIN VH GENE
TITLE OF INVENTION: SEGMENTS AND DNA FRAGMENTS CONTAINING THE SAME
NUMBER OF SEQUENCES: 145
CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson, P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02110-2804

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible

OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/545,809A
FILING DATE: 27-MAR-1996

PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP93/00603
FILING DATE: 10-MAY-1993

ATTORNEY/AGENT INFORMATION:
NAME: Freeman, John W.
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 06501/004001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-542-5070

TELEFAX: 617-542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 99:
SEQUENCE CHARACTERISTICS:
LENGTH: 117 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-545-809A-99

Query Match 67.1%; Score 57; DB 3; Length 117;
Best Local Similarity 75.0%; Pred. No. 0.037;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVKG 17
||:|||||
Db 70 ISSSGSTIYADSVKG 85

RESULT 14
US-08-665-202-30
; Sequence 30, Application US/08665202
; Patent No. 5977322
; GENERAL INFORMATION:
; APPLICANT: Marks, James D.
; APPLICANT: Schier, Robert
; TITLE OF INVENTION: No. 5977322el High Affinity Human Antibodies to
; TITLE OF INVENTION: Tumor Antigens
; NUMBER OF SEQUENCES: 141
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/665,202
; FILING DATE: 13-JUN-1996
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/000,238
; FILING DATE: 14-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/000,250
; FILING DATE: 15-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Hunter, Tom
; REGISTRATION NUMBER: 38,498
; REFERENCE/DOCKET NUMBER: 02307E-061410
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 123 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-665-202-30

Query Match 67.1%; Score 57; DB 2; Length 123;
Best Local Similarity 75.0%; Pred. No. 0.04;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVKG 17
||:|||||

Db 51 ISSSGSTIYADSVKG 66

RESULT 15

US-09-315-574-30
; Sequence 30, Application US/09315574
; Patent No. 6512097
; GENERAL INFORMATION:
; APPLICANT: Marks, James D.

; APPLICANT: Schier, Robert
; TITLE OF INVENTION: No. 6512097el High Affinity Human Antibodies to
; TITLE OF INVENTION: Tumor Antigens
; NUMBER OF SEQUENCES: 141
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Majestic, Parsons, Siebert & Hsue P.C.
; STREET: Four Embarcadero Center, Suite 1100
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4106

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/315,574
; FILING DATE: 20-MAY-99
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/000,238
; FILING DATE: 14-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/000,250
; FILING DATE: 15-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/665,202
; FILING DATE: 13-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hunter, Tom
; REGISTRATION NUMBER: 38,498
; REFERENCE/DOCKET NUMBER: 02307E-061411
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 123 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-315-574-30

Query Match 67.1%; Score 57; DB 4; Length 123;
Best Local Similarity 75.0%; Pred. No. 0.04;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVKG 17
||:|||||

Db 51 ISSSGSTIYADSVKG 66

Search completed: April 28, 2005, 18:28:08
Job time : 34.5312 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2005, 19:24:46 ; Search time 105.453 Seconds
(without alignments)
53.700 Million cell updates/sec

Title: US-10-088-639A-2_COPY_177_193

Perfect score: 85
Sequence: 1 VISENGRTINYADSVKG 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1426032 seqs, 333106140 residues

Total number of hits satisfying chosen parameters: 1426032

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	65	76.5	17	14 US-10-031-874A-172	Sequence 172, App
2	61	71.8	248	15 US-09-880-748-1648	Sequence 1648, Ap
3	61	71.8	248	15 US-10-293-418-1648	Sequence 1648, Ap
4	59	69.4	129	10 US-09-848-798-143	Sequence 143, App
5	58	68.2	120	17 US-10-492-668-105	Sequence 105, App
6	58	68.2	120	17 US-10-492-668-109	Sequence 109, App
7	58	68.2	261	15 US-09-880-748-1503	Sequence 1503, Ap
8	58	68.2	261	15 US-10-293-418-1503	Sequence 1503, Ap
9	57	67.1	98	13 US-10-066-895-4	Sequence 4, Appli
10	57	67.1	98	14 US-10-194-975-17	Sequence 17, Appl
11	57	67.1	98	15 US-10-308-817-57	Sequence 57, Appl
12	57	67.1	98	15 US-10-032-037B-63	Sequence 63, Appl
13	57	67.1	98	15 US-10-032-037B-79	Sequence 79, Appl

14	67.1	98	15	US-10-029-988B-63	Sequence 63, Appl
15	67.1	98	15	US-10-029-988B-79	Sequence 79, Appl
16	67.1	98	15	US-10-032-423A-63	Sequence 63, Appl
17	67.1	98	15	US-10-032-423A-79	Sequence 79, Appl
18	67.1	98	15	US-10-453-698-57	Sequence 57, Appl
19	67.1	98	15	US-10-029-926B-63	Sequence 63, Appl
20	67.1	98	15	US-10-029-926B-79	Sequence 79, Appl
21	67.1	98	15	US-10-038-591-30	Sequence 30, Appl
22	67.1	98	16	US-10-379-392-16	Sequence 16, Appl
23	67.1	98	16	US-10-884-830-601	Sequence 601, App
24	67.1	98	16	US-10-884-830-639	Sequence 639, App
25	67.1	98	17	US-10-887-954-4	Sequence 4, Appli
26	67.1	109	15	US-10-309-764-23	Sequence 23, Appl
27	67.1	109	17	US-10-727-155-288	Sequence 288, App
28	67.1	114	17	US-10-938-353-115	Sequence 115, App
29	67.1	115	17	US-10-938-353-108	Sequence 108, App
30	67.1	115	17	US-10-938-353-116	Sequence 116, App
31	67.1	116	17	US-10-727-155-84	Sequence 84, Appl
32	67.1	116	17	US-10-727-155-140	Sequence 140, App
33	67.1	116	17	US-10-727-155-148	Sequence 148, App
34	67.1	117	17	US-10-938-353-106	Sequence 106, App
35	67.1	122	14	US-10-269-805-7	Sequence 7, Appli
36	67.1	122	14	US-10-269-805-11	Sequence 11, Appl
37	67.1	122	14	US-10-269-805-53	Sequence 53, Appl
38	67.1	122	17	US-10-888-959-5	Sequence 5, Appli
39	67.1	131	10	US-09-988-115A-53	Sequence 53, Appl
40	67.1	236	10	US-09-880-748-2010	Sequence 2010, Ap
41	67.1	236	15	US-10-293-418-2010	Sequence 2010, Ap
42	67.1	237	10	US-09-880-748-2003	Sequence 2003, Ap
43	67.1	237	10	US-09-880-748-2005	Sequence 2005, Ap
44	67.1	237	10	US-09-880-748-2006	Sequence 2006, Ap
45	67.1	237	10	US-09-880-748-2017	Sequence 2017, Ap

ALIGNMENTS

RESULT 1

US-10-031-874A-172
; Sequence 172, Application US/10031874A
; Publication No. US20030190598A1
; GENERAL INFORMATION:
; APPLICANT: TANHA, JAMSHID
; APPLICANT: DUBUC, GINETTE
; APPLICANT: NARANG, SARAN
; TITLE OF INVENTION: SINGLES-DOMAIN ANTIGEN-BINDING ANTIBODY FRAGMENTS
; TITLE OF INVENTION: DERIVED FROM LLAMA ANTIBODIES
; FILE REFERENCE: 11054-1
; CURRENT APPLICATION NUMBER: US/10/031,874A
; CURRENT FILING DATE: 2002-11-14
; PRIOR APPLICATION NUMBER: 60/207,234
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 212
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 172
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Lama glama
US-10-031-874A-172

Query Match 76.5%; Score 65; DB 14; Length 17;
Best Local Similarity 70.6%; Pred. No. 0.00067;
Matches 12; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 VISENGRTINYADSVKG 17
::: ||| |||||
Db 1 LLRSRGRTTNYADSVKG 17

RESULT 2

US-09-880-748-1648
; Sequence 1648, Application US/09880748
; Publication No. US20030059937A1

RESULT 6
 US-10-492-668-109
 ; Sequence 109, Application US/10492668
 ; Publication No. US20050054001A1
 ; GENERAL INFORMATION:
 ; APPLICANT: VLAAMS INTERUNIVERSITAIR INSTITUUT VOOR BIOTECHNOLOGIE VZW
 ; TITLE OF INVENTION: FUNCTIONAL HEAVY CHAIN ANTIBODIES, FRAGMENTS THEREOF, LIBRARY THEREOF
 ; FILE REFERENCE: VIB-030-PCT
 ; CURRENT APPLICATION NUMBER: US/10/492,668
 ; CURRENT FILING DATE: 2004-04-15
 ; PRIOR APPLICATION NUMBER: EP01204037.4
 ; PRIOR FILING DATE: 2001-10-24
 ; PRIOR APPLICATION NUMBER: JP2002-004184
 ; PRIOR FILING DATE: 2002-01-11
 ; PRIOR APPLICATION NUMBER: US60/335,054
 ; PRIOR FILING DATE: 2001-10-24
 ; NUMBER OF SEQ ID NOS: 165
 ; SOFTWARE: Patent in version 3.1
 ; SEQ ID NO 109
 ; LENGTH: 120
 ; TYPE: PRT
 ; ORGANISM: Camelus dromedarius
 ; NAME/KEY: MISC FEATURE
 ; OTHER INFORMATION: C24-A
 US-10-492-668-109

Query Match 68.2%; Score 58; DB 17; Length 120;
 Best Local Similarity 70.8%; Pred. No. 0.09;
 Matches 12; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 VISENGRTINYADSVKG 17
 |||:|||||
 Db 50 VISNDGRYDYADSVKG 66

RESULT 7
 US-09-880-748-1503
 ; Sequence 1503, Application US/09880748
 ; Publication No. US20030059937A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ruben et al.
 ; TITLE OF INVENTION: Antibodies that Immunosepecifically Bind Blys
 ; FILE REFERENCE: PF523
 ; CURRENT APPLICATION NUMBER: US/09/880,748
 ; CURRENT FILING DATE: 2001-06-15
 ; PRIOR APPLICATION NUMBER: 60/212,210
 ; PRIOR FILING DATE: 2000-06-15
 ; PRIOR APPLICATION NUMBER: 60/240,816
 ; PRIOR FILING DATE: 2000-10-17
 ; PRIOR APPLICATION NUMBER: 60/276,248
 ; PRIOR FILING DATE: 2001-03-16
 ; PRIOR APPLICATION NUMBER: 60/277,379
 ; PRIOR FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: 60/293,499
 ; PRIOR FILING DATE: 2001-05-25
 ; NUMBER OF SEQ ID NOS: 3239
 ; SOFTWARE: Patent in Ver. 2.0
 ; SEQ ID NO 1503
 ; LENGTH: 261
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-880-748-1503

Query Match 68.2%; Score 58; DB 10; Length 261;
 Best Local Similarity 76.5%; Pred. No. 0.22;
 Matches 13; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 VISENGRTINYADSVKG 17
 |||:|||||
 Db 52 VISYNGRTKYLDYADSVKG 68

RESULT 8
 US-10-293-418-1503
 ; Sequence 1503, Application US/10293418
 ; Publication No. US20030223996A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ruben et al.
 ; TITLE OF INVENTION: Antibodies that Immunosepecifically Bind Blys
 ; FILE REFERENCE: PF523P2
 ; CURRENT APPLICATION NUMBER: US/10/293,418
 ; CURRENT FILING DATE: 2002-11-27
 ; PRIOR APPLICATION NUMBER: 60/331,469
 ; PRIOR FILING DATE: 2001-11-16
 ; PRIOR APPLICATION NUMBER: 60/340,817
 ; PRIOR FILING DATE: 2001-12-19
 ; PRIOR APPLICATION NUMBER: 09/880,748
 ; PRIOR FILING DATE: 2001-06-15
 ; PRIOR APPLICATION NUMBER: 60/293,499
 ; PRIOR FILING DATE: 2001-05-25
 ; PRIOR APPLICATION NUMBER: 60/277,379
 ; PRIOR FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: 60/276,248
 ; PRIOR FILING DATE: 2001-03-16
 ; PRIOR APPLICATION NUMBER: 60/240,816
 ; PRIOR FILING DATE: 2000-10-17
 ; PRIOR APPLICATION NUMBER: 60/212,210
 ; PRIOR FILING DATE: 2000-06-16
 ; NUMBER OF SEQ ID NOS: 3247
 ; SEQ ID NO 1503
 ; LENGTH: 261
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-293-418-1503

Query Match 68.2%; Score 58; DB 15; Length 261;
 Best Local Similarity 76.5%; Pred. No. 0.22;
 Matches 13; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 VISENGRTINYADSVKG 17
 |||:|||||
 Db 52 VISYNGRTKYLDYADSVKG 68

RESULT 9
 US-10-066-895-4
 ; Sequence 4, Application US/10066895
 ; Publication No. US20020141990A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Deen, Keith C.
 ; DILLON, Susan B.
 ; PORTER, Terence C.
 ; SWEET, Raymond A.
 ; TITLE OF INVENTION: Human Monoclonal Antibodies
 ; NUMBER OF SEQUENCES: 30
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSER: SmithKline Beecham Corporation
 ; STREET: 709 Swedeland Road
 ; CITY: King of Prussia
 ; STATE: PA
 ; COUNTRY: U.S.A.
 ; ZIP: 19046
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/10/066,895
 ; FILING DATE: 04-Feb-2002
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 09/297,344
 ; FILING DATE: 1999-JUN-09

```
;
; APPLICATION NUMBER: 60/030,149
; FILING DATE: 01-NOV-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Geiger, Kathleen
; REGISTRATION NUMBER: 35,880
; REFERENCE/DOCKET NUMBER: P50504
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5968
; TELEFAX: 610-270-5090
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 98 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-066-895-4
Query Match 67.1%; Score 57; DB 13; Length 98;
Best Local Similarity 75.0%; Pred. No. 0.1;
Matches 12; Conservative 1; Mismatches 3; Indels 3; Gaps 0;

QY 2 ISENGRTINYADSVKG 17
Db 51 ISSSGSTIYADSVKG 66

RESULT 10
US-10-194-975-17
; Sequence 17, Application US/10194975
; Publication No. US20030039649A1
; GENERAL INFORMATION:
; APPLICANT: Foote, Jefferson
; TITLE OF INVENTION: Super Humanized Antibodies
; FILE REFERENCE: 501231.01
; CURRENT APPLICATION NUMBER: US/10/194,975
; CURRENT FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: US 60/305,111
; PRIOR FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-194-975-17
Query Match 67.1%; Score 57; DB 14; Length 98;
Best Local Similarity 75.0%; Pred. No. 0.1;
Matches 12; Conservative 1; Mismatches 3; Indels 3; Gaps 0;

QY 2 ISENGRTINYADSVKG 17
Db 51 ISSSGSTIYADSVKG 66

RESULT 11
US-10-308-817-57
; Sequence 57, Application US/10308817
; Publication No. US20030219861A1
; GENERAL INFORMATION:
; APPLICANT: Rother, Russell
; APPLICANT: Wu, Dayang
; TITLE OF INVENTION: HYBRID ANTIBODIES
; FILE REFERENCE: 1087-37
; CURRENT APPLICATION NUMBER: US/10/308,817
; CURRENT FILING DATE: 2002-12-03
; NUMBER OF SEQ ID NOS: 195
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 57
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-308-817-57
Query Match 67.1%; Score 57; DB 15; Length 98;
Best Local Similarity 75.0%; Pred. No. 0.1;
Matches 12; Conservative 1; Mismatches 3; Indels 3; Gaps 0;

QY 2 ISENGRTINYADSVKG 17
Db 51 ISSSGSTIYADSVKG 66

RESULT 12
US-10-032-037B-63
; Sequence 63, Application US/10032037B
; Publication No. US20040001822A1
; GENERAL INFORMATION:
; APPLICANT: Bio-Technology General Corp.
; TITLE OF INVENTION: Y17-ISOLATED MOLECULES COMPRISING EPITOPES CONTAINING SULFATED
; FILE REFERENCE: 10793/44
; CURRENT APPLICATION NUMBER: US/10/032,037B
; CURRENT FILING DATE: 2001-12-31
; PRIOR APPLICATION NUMBER: 60/258,948
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 204
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 63
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-032-037B-63
Query Match 67.1%; Score 57; DB 15; Length 98;
Best Local Similarity 75.0%; Pred. No. 0.1;
Matches 12; Conservative 1; Mismatches 3; Indels 3; Gaps 0;

QY 2 ISENGRTINYADSVKG 17
Db 51 ISSSGSTIYADSVKG 66

RESULT 13
US-10-032-037B-79
; Sequence 79, Application US/10032037B
; Publication No. US20040001822A1
; GENERAL INFORMATION:
; APPLICANT: Bio-Technology General Corp.
; TITLE OF INVENTION: Y17-ISOLATED MOLECULES COMPRISING EPITOPES CONTAINING SULFATED
; FILE REFERENCE: 10793/44
; CURRENT APPLICATION NUMBER: US/10/032,037B
; CURRENT FILING DATE: 2001-12-31
; PRIOR APPLICATION NUMBER: 60/258,948
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 204
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 79
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-032-037B-79
Query Match 67.1%; Score 57; DB 15; Length 98;
Best Local Similarity 75.0%; Pred. No. 0.1;
Matches 12; Conservative 1; Mismatches 3; Indels 3; Gaps 0;

QY 2 ISENGRTINYADSVKG 17
Db 51 ISSSGSTIYADSVKG 66
```

RESULT 14
US-10-029-988B-63
; Sequence 63, Application US/10029988B
; Publication No. US20040001839A1
; GENERAL INFORMATION:
; APPLICANT: Bio-technology General Corp.
; TITLE OF INVENTION: Y17-ISOLATED MOLECULES COMPRISING EPITOPES CONTAINING SULFATED
; TITLE OF INVENTION: MOIETIES, ANTIBODIES TO SUCH EPITOPES, AND USES THEREOF
; FILE REFERENCE: 10793/46
; CURRENT APPLICATION NUMBER: US/10/029,988B
; CURRENT FILING DATE: 2001-12-31
; PRIOR APPLICATION NUMBER: 60/258,948
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 204
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 63
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-029-988B-63

Query Match 67.1%; Score 57; DB 15; Length 98;
Best Local Similarity 75.0%; Pred. No. 0.1;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVKG 17
||:|||||
Db 51 ISSSGSTIYYADSVKG 66

RESULT 15
US-10-029-988B-79
; Sequence 79, Application US/10029988B
; Publication No. US20040001839A1
; GENERAL INFORMATION:
; APPLICANT: Bio-technology General Corp.
; TITLE OF INVENTION: Y17-ISOLATED MOLECULES COMPRISING EPITOPES CONTAINING SULFATED
; TITLE OF INVENTION: MOIETIES, ANTIBODIES TO SUCH EPITOPES, AND USES THEREOF
; FILE REFERENCE: 10793/46
; CURRENT APPLICATION NUMBER: US/10/029,988B
; CURRENT FILING DATE: 2001-12-31
; PRIOR APPLICATION NUMBER: 60/258,948
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 204
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 79
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-029-988B-79

Query Match 67.1%; Score 57; DB 15; Length 98;
Best Local Similarity 75.0%; Pred. No. 0.1;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVKG 17
||:|||||
Db 51 ISSSGSTIYYADSVKG 66

Search completed: April 28, 2005, 19:02:16
Job time : 105.453 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2005, 18:05:55 ; Search time 23.375 Seconds
(without alignments)
69.976 Million cell updates/sec

Title: US-10-088-639A-2_COPY_177_193

Perfect score: 85

Sequence: 1 VISENGRTINYADSVKG 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	58	68.2	115	2 S57445	Ig heavy chain V-J
2	57	67.1	97	2 S24252	Ig heavy chain V r
3	57	67.1	98	2 S26891	Ig heavy chain V r
4	57	67.1	98	2 S26930	Ig heavy chain V r
5	57	67.1	110	2 PH1652	Ig heavy chain V r
6	57	67.1	114	2 S31120	Ig heavy chain - h
7	57	67.1	118	2 S31105	Ig heavy chain (eu
8	57	67.1	128	2 S26790	Ig heavy chain V r
9	57	67.1	130	2 I37783	Ig variable region
10	57	67.1	132	2 I47191	Ig heavy chain var
11	56	65.9	97	2 PH0875	Ig heavy chain V r
12	53	62.4	120	2 S36273	Ig heavy chain V r
13	52	61.2	108	2 S24251	Ig heavy chain V r
14	52	61.2	115	2 S36267	Ig heavy chain V r
15	52	61.2	123	2 S30532	Ig heavy chain V r
16	51	60.0	98	2 S26894	Ig heavy chain V r
17	51	60.0	110	2 S36282	Ig heavy chain V r
18	51	60.0	114	1 M3HWE	Ig heavy chain V-I
19	51	60.0	117	2 S21980	Ig heavy chain V-g
20	51	60.0	121	2 PH1661	Ig heavy chain V r
21	51	60.0	123	2 S26794	Ig heavy chain V r
22	51	60.0	125	2 S30531	Ig heavy chain V r
23	51	60.0	143	2 S23624	Ig heavy chain V r
24	50	58.8	98	2 PH0874	Ig heavy chain V r
25	50	58.8	98	2 S26934	Ig heavy chain V r
26	50	58.8	98	2 S26927	Ig heavy chain V r
27	50	58.8	113	2 S57441	Ig heavy chain V-J
28	50	58.8	121	2 S31104	Ig heavy chain (eu
29	50	58.8	128	2 S31595	Ig heavy chain V r

30	50	58.8	133	2 I47195	Ig heavy chain var
31	50	58.8	151	2 A60943	Ig heavy chain pre
32	50	58.8	191	2 JL0048	Ig heavy chain V r
33	49	57.6	95	2 S20777	Ig heavy chain V r
34	49	57.6	120	1 G1HUD8	Ig heavy chain V-I
35	49	57.6	120	2 S49590	Ig heavy chain V r
36	49	57.6	121	2 B34871	Ig heavy chain V r
37	49	57.6	122	1 A1HUTR	Ig heavy chain V-I
38	49	57.6	123	2 PC4281	anti-SS-A/Ro 60K p
39	48	56.5	94	2 D25913	Ig heavy chain V r
40	48	56.5	96	2 S20781	Ig heavy chain V r
41	48	56.5	98	2 S29545	Ig heavy chain V r
42	48	56.5	98	2 S26928	Ig heavy chain V r
43	48	56.5	118	2 S00700	Ig heavy chain V r
44	48	56.5	135	2 I37778	Ig variable region
45	48	56.5	136	1 G1MS21	Ig heavy chain pre

ALIGNMENTS

RESULT 1

S57445

Ig heavy chain V-J region - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 10-Oct-1995 #sequence_revision 17-Nov-1995 #text_change 23-Jul-1999

C:Accession: S57445

R:Paterson, G.; Wilson, G.; Kennedy, P.G.E.; Willison, H.J.

Submitted to the EMBL Data Library, June 1995

A:Description: Analysis of anti-GM1 ganglioside IgM antibodies cloned from motor neuropat

A:Reference number: S57408

A:Accession: S57445

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-115 <PAT>

A:Cross-references: EMBL:X87891; NID:g871285; PIDN:CAA61142.1; PID:g871286

C:Genetics:

A:Introns: 99/2

C:Superfamily: immunoglobulin V region; immunoglobulin homology

F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 68.2%; Score 58; DB 2; Length 115;
Best Local Similarity 76.5%; Pred. No. 0.04;
Matches 13; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY	1	VISENGRTINYADSVKG	17
DB	50	VISANGRDYKYYADSVKG	66

RESULT 2

S24252

Ig heavy chain V region (N54P3-D-JH6) - human

C:Species: Homo sapiens (man)

C>Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999

C:Accession: S24252

R:Stewart, A.K.; Huang, C.; Stollar, B.D.; Schwartz, R.S.

Submitted to the EMBL Data Library, June 1992

A:Description: A single VH gene predominates in the rearranged and expressed human B cell

A:Reference number: S24247

A:Accession: S24252

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-97 <STE>

A:Cross-references: EMBL:X67073; NID:g38403; PIDN:CAA47458.1; PID:g38404

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

Query Match 67.1%; Score 57; DB 2; Length 97;
Best Local Similarity 75.0%; Pred. No. 0.049;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY	2	ISENGRTINYADSVKG	17
----	---	------------------	----

```
Db      28 ISSSGSTIYADSVKG 43
      ||.:|||
RESULT 3
S26891
Ig heavy chain V region (DP-58) - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999
C:Accession: S26891
R:Tomlinson, I.M.; Walter, G.; Marks, J.D.; Llewellyn, M.B.; Winter, G.
J. Mol. Biol. 227, 776-798, 1992
A:Title: The repertoire of human germline V(H) sequences reveals about fifty groups of
A:Reference number: S26885; MUID:93021117; PMID:1404388
A:Accession: S26891
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-98 <TOM>
A:Cross-references: EMBL:Z12358; NID:g32935; PIDN:CAA78228.1; PID:g32936
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMM>
      Query Match      67.1%; Score 57; DB 2; Length 98;
      Best Local Similarity 75.0%; Pred. No. 0.049;
      Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
      QY      2 ISENGRTIYADSVKG 17
      Db      51 ISSSGSTIYADSVKG 66
      ||.:|||
RESULT 4
S26930
Ig heavy chain V region (DP-35) - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999
C:Accession: S26930
R:Tomlinson, I.M.; Walter, G.; Marks, J.D.; Llewellyn, M.B.; Winter, G.
J. Mol. Biol. 227, 776-798, 1992
A:Title: The repertoire of human germline V(H) sequences reveals about fifty groups of
A:Reference number: S26885; MUID:93021117; PMID:1404388
A:Accession: S26930
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-98 <TOM>
A:Cross-references: EMBL:Z12337; NID:g32892; PIDN:CAA78207.1; PID:g32893
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMM>
      Query Match      67.1%; Score 57; DB 2; Length 98;
      Best Local Similarity 75.0%; Pred. No. 0.049;
      Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
      QY      2 ISENGRTIYADSVKG 17
      Db      51 ISSSGSTIYADSVKG 66
      ||.:|||
RESULT 5
PH1652
Ig heavy chain V region (clone 5B4) - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 16-Aug-1996
C:Accession: PH1652
R:Hillson, J.L.; Karr, N.S.; Opplinger, I.R.; Mannik, M.; Sasso, E.H.
J. Exp. Med. 178, 331-336, 1993
A:Title: The structural basis of germline-encoded VH3 immunoglobulin binding to staphylo
A:Reference number: PH1642; MUID:93301610; PMID:8315388
A:Accession: PH1652
A:Molecule type: mRNA
A:Residues: 1-110 <HIL>
```

```
A:Experimental source: B cell
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:7-90/Domain: immunoglobulin homology <IMM>
      Query Match      67.1%; Score 57; DB 2; Length 110;
      Best Local Similarity 75.0%; Pred. No. 0.055;
      Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
      QY      2 ISENGRTIYADSVKG 17
      Db      43 ISSSGSTIYADSVKG 58
      ||.:|||
RESULT 6
S31120
Ig heavy chain - human
C:Species: Homo sapiens (man)
C>Date: 02-Dec-1993 #sequence_revision 26-May-1995 #text_change 17-Mar-1999
C:Accession: S31120
R:Raaphorst, F.M.; Timmers, E.; Kenter, M.J.H.; van Tol, M.J.D.; Vossen, J.M.; Schuurman,
Eur. J. Immunol. 22, 247-251, 1992
A:Title: Restricted utilization of germ-line V(H)3 genes and short diverse third comple
A:Reference number: S31104; MUID:92111633; PMID:1730252
A:Accession: S31120
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: mRNA
A:Residues: 1-114 <RAA>
A:Cross-references: EMBL:X62972
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1991
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMM>
      Query Match      67.1%; Score 57; DB 2; Length 114;
      Best Local Similarity 75.0%; Pred. No. 0.057;
      Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
      QY      2 ISENGRTIYADSVKG 17
      Db      51 ISSSGSTIYADSVKG 66
      ||.:|||
RESULT 7
S31105
Ig heavy chain (subclass IgM) - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 22-Nov-1993 #sequence_revision 26-May-1995 #text_change 23-Jul-1999
C:Accession: S31105
R:Raaphorst, F.M.; Timmers, E.; Kenter, M.J.H.; van Tol, M.J.D.; Vossen, J.M.; Schuurman,
Eur. J. Immunol. 22, 247-251, 1992
A:Title: Restricted utilization of germ-line V(H)3 genes and short diverse third comple
A:Reference number: S31104; MUID:92111633; PMID:1730252
A:Accession: S31105
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: mRNA
A:Residues: 1-118 <RAA>
A:Cross-references: EMBL:X63081; NID:g32648; PIDN:CAA44803.1; PID:g32649
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1991
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMM>
      Query Match      67.1%; Score 57; DB 2; Length 118;
      Best Local Similarity 75.0%; Pred. No. 0.06;
      Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
      QY      2 ISENGRTIYADSVKG 17
      Db      51 ISSSGSTIYADSVKG 66
      ||.:|||
RESULT 8
PH1652
Ig heavy chain V region (clone 5B4) - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 16-Aug-1996
C:Accession: PH1652
R:Hillson, J.L.; Karr, N.S.; Opplinger, I.R.; Mannik, M.; Sasso, E.H.
J. Exp. Med. 178, 331-336, 1993
A:Title: The structural basis of germline-encoded VH3 immunoglobulin binding to staphylo
A:Reference number: PH1642; MUID:93301610; PMID:8315388
A:Accession: PH1652
A:Molecule type: mRNA
A:Residues: 1-110 <HIL>
```

S26790
Ig heavy chain V region - human
C:Species: Homo sapiens (man)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 20-Jun-2000
C:Accession: S26790
R:Mortari, P.; Newton, J.A.; Wang, J.Y.; Schroeder Jr., H.W.
Eur. J. Immunol. 22, 241-245, 1992
A:Title: The human cord blood antibody repertoire. Frequent usage of the V(H)7 gene family.
A:Reference number: S26786; MUID:92111632; PMID:1730251
A:Accession: S26790
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-128 <NR>
A:Cross-references: EMBL:X61013; NID:G32798; PIDN:CAA43347.1; PID:G1335128
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMM>
Query Match 67.1%; Score 57; DB 2; Length 128;
Best Local Similarity 75.0%; Pred. No. 0.065;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 2 ISENGRTINYADSVK 17
DB 51 ISSSGRTIYADSVK 66
RESULT 9
137783
Ig variable region (VDJ) (clone T24-3) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 16-Feb-1996 #sequence_revision 16-Feb-1996 #text_change 23-Jul-1999
C:Accession: I37783; S25477
R:Demaison, C.; Chastagner, P.; Theze, J.; Zouali, M.
Proc. Natl. Acad. Sci. U.S.A. 91, 514-518, 1994
A:Title: Somatic diversification in the heavy chain variable region genes expressed by B cells.
A:Reference number: A36876; MUID:94119917; PMID:8290556
A:Accession: I37783
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-130 <RES>
A:Cross-references: EMBL:X67907; NID:G33584; PIDN:CAA48105.1; PID:G33585
C:Superfamily: immunoglobulin V region; immunoglobulin homology
F:22-105/Domain: immunoglobulin homology <IMM>
Query Match 67.1%; Score 57; DB 2; Length 130;
Best Local Similarity 68.8%; Pred. No. 0.066;
Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 2 ISENGRTINYADSVK 17
DB 58 MSSSGRTIYADSVK 73
RESULT 10
147191
Ig heavy chain variable VDJ region - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Jan-2000
C:Accession: I47191
R:Sun, J.; Kacsokovics, I.; Brown, W.R.; Butler, J.E.
J. Immunol. 153, 5618-5627, 1994
A:Title: Expressed swine VH genes belong to a small VH gene family homologous to human V genes.
A:Reference number: I47177; MUID:95081609; PMID:7989761
A:Accession: I47191
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-132 <SUN>
A:Cross-references: EMBL:U15450; NID:G571386; PIDN:AAA67016.1; PID:G571387
C:Superfamily: immunoglobulin V region; immunoglobulin homology
F:34-117/Domain: immunoglobulin homology <IMM>
Query Match 67.1%; Score 57; DB 2; Length 132;

Best Local Similarity 68.8%; Pred. No. 0.067;
Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 2 ISENGRTINYADSVK 17
DB 70 ISDSGRTDYADSVK 85
RESULT 11
PH0875
Ig heavy chain V region (anti-DNA, H2F) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 09-Jul-2004
C:Accession: PH0875
R:Manheimer-Lory, A.; Katz, J.B.; Pillinger, M.; Ghossein, C.; Smith, A.; Diamond, B.
J. Exp. Med. 174, 1639-1652, 1991
A:Title: Molecular characteristics of antibodies bearing an anti-DNA-associated idiotype.
A:Reference number: PH0862; MUID:92078875; PMID:1660528
A:Accession: PH0875
A:Molecule type: DNA
A:Residues: 1-97 <MAN>
A:Cross-references: UNIPROT:Q9UL91
C:Comment: This antibody is produced by Epstein-Barr virus-transformed B cell that bears
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:14-97/Domain: immunoglobulin homology <IMM>
F:30-35/Region: complementarity-determining 1
F:49-66/Region: complementarity-determining 2
Query Match 65.9%; Score 56; DB 2; Length 97;
Best Local Similarity 75.0%; Pred. No. 0.071;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 ISENGRTINYADSVK 17
DB 50 ISSSGRTIYADSVK 65
RESULT 12
S36273
Ig heavy chain V region (clone alpha-THY-32) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 23-Jul-1999
C:Accession: S36273
R:Griffiths, A.D.; Malmqvist, M.; Marks, J.D.; Bye, J.M.; McCafferty, J.; WEMO J. 12, 725-734, 1993
A:Title: Human anti-self antibodies with high specificity from phage display libraries.
A:Reference number: S36256; MUID:93178448; PMID:7679990
A:Accession: S36273
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-120 <GRI>
A:Cross-references: EMBL:Z18834; NID:G33116; PIDN:CAA79286.1; PID:G939896
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMM>
Query Match 62.4%; Score 53; DB 2; Length 120;
Best Local Similarity 68.8%; Pred. No. 0.27;
Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 2 ISENGRTINYADSVK 17
DB 51 ISSNSGSGIYADSVK 66
RESULT 13
S24251
Ig heavy chain V region (N54P3-D-JH4) - human
C:Species: Homo sapiens (man)
C:Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999
C:Accession: S24251
R:Stewart, A.K.; Huang, C.; Stollar, B.D.; Schwartz, R.S.

submitted to the EMBL Data Library, June 1992
A;Description: A single VH gene predominates in the rearranged and expressed human B cell
A;Reference number: S24247
A;Accession: S24251
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-108 <STE>
A;Cross-references: EMBL:X67072; NID:g38401; PIDN:CAA47457.1; PID:g38402
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotrimer; immunoglobulin
F;11-94/Domain: immunoglobulin homology <IMM>

Query Match 61.2%; Score 52; DB 2; Length 108;
Best Local Similarity 68.8%; Pred. No. 0.35;
Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVKG 17
||| : ||| |||
Db 47 ISSGSAIYYADSVKG 62

RESULT 14
S36267
Ig heavy chain V region (clone alpha-THY-33) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 23-Jul-1999
C;Accession: S36267
R;Griffiths, A.D.; Malmqvist, M.; Marks, J.D.; Bye, J.M.; Embleton, M.J.; McCafferty, J.
EMBO J. 12, 725-734, 1993
A;Title: Human anti-self antibodies with high specificity from phage display libraries.
A;Reference number: S36256; MUID:93178448; PMID:7679990
A;Accession: S36267
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-115 <GR1>
A;Cross-references: EMBL:Z18836; NID:g33117; PIDN:CAA79288.1; PID:g939897
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotrimer; immunoglobulin
F;15-98/Domain: immunoglobulin homology <IMM>

Query Match 61.2%; Score 52; DB 2; Length 115;
Best Local Similarity 62.5%; Pred. No. 0.37;
Matches 10; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVKG 17
||| : ||| |||
Db 51 ISYDGRSVYYADSVQ 66

RESULT 15
S30532
Ig heavy chain V region - human
C;Species: Homo sapiens (man)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C;Accession: S30532
R;Mariette, X.
submitted to the EMBL Data Library, October 1992
A;Reference number: S30520
A;Accession: S30532
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-123 <VAR>
A;Cross-references: UNIPROT:Q8WU38; EMBL:Z18318
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotrimer; immunoglobulin
F;15-98/Domain: immunoglobulin homology <IMM>

Query Match 61.2%; Score 52; DB 2; Length 123;
Best Local Similarity 68.8%; Pred. No. 0.4;
Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVKG 17
||| : ||| |||

Db 51 ISWNSGTLGYADSVKG 66

Search completed: April 28, 2005, 18:26:03
Job time : 30.375 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 01:42:46 ; Search time 14.1132 Seconds
(without alignments)
62.721 Million cell updates/sec

Title: US-10-088-639A-2_COPY_177_193

Perfect score: 85

Sequence: 1 VISENGRTINYADSVKG 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	51	60.0	114	1	HV3B HUMAN
2	49	57.6	120	1	HV3U HUMAN
3	49	57.6	122	1	HV3A HUMAN
4	48	56.5	136	1	HV16 MOUSE
5	47	55.3	116	1	HV3T HUMAN
6	46	54.1	117	1	HV03 CARAU
7	46	54.1	122	1	HV3G HUMAN
8	45	52.9	119	1	HV3M HUMAN
9	45	52.9	119	1	HV3N HUMAN
10	45	52.9	126	1	HV3K HUMAN
11	44	51.8	268	1	FLHP BACSU
12	43	50.6	117	1	HV03 CAICR
13	43	50.6	119	1	HV3I HUMAN
14	43	50.6	383	1	GALI YERPE
15	43	50.6	502	1	NADB NEIMA
16	43	50.6	502	1	NADB NEIMB
17	43	50.6	658	1	UVRB LISIN
18	43	50.6	658	1	UVRB LISMO
19	43	50.6	1156	1	PHYB SOYBN
20	42.5	50.0	117	1	HV02 CANFA
21	42	49.4	115	1	HV3F HUMAN
22	42	49.4	121	1	HV3J HUMAN
23	42	49.4	201	1	HAMI BORBU
24	42	49.4	341	1	MBNL_MOUSE
25	42	49.4	388	1	MBNL HUMAN
26	42	49.4	582	1	YMA5 CAEEL
27	42	49.4	659	1	SL54 HUMAN
28	41.5	48.8	718	1	SL53 HUMAN
29	41	48.2	98	1	HV57_MOUSE
30	41	48.2	117	1	HV3C HUMAN
31	41	48.2	117	1	HV53_MOUSE
32	41	48.2	119	1	HV3L HUMAN
33	41	48.2	260	1	RS4_YEAST

34	41	48.2	262	1	RS4 CANAL
35	41	48.2	342	1	MBN3_MOUSE
36	41	48.2	354	1	MBN3_HUMAN
37	41	48.2	473	1	GLGS_WHEAT
38	41	48.2	489	1	GLGS_BETVU
39	41	48.2	508	1	GLGS_VICFA
40	41	48.2	513	1	GLGS_HORVU
41	41	48.2	520	1	GLGS_ARATH
42	41	48.2	566	1	HEMA_TAZIN
43	41	48.2	1451	1	MYM1_HUMAN
44	40.5	47.6	116	1	HV05_CARAU
45	40	47.1	122	1	HV3H_HUMAN
46	40	47.1	284	1	AROE_STRPN
47	40	47.1	435	1	YYS3_CAEEL
48	40	47.1	483	1	GLGS_ORYSA
49	40	47.1	508	1	CPT7_CAVPO
50	40	47.1	890	1	MUIS_RICCN
51	40	47.1	1044	1	SYI_METTM
52	39	45.9	117	1	HV42_MOUSE
53	39	45.9	238	1	Y437_STAEP
54	39	45.9	238	1	Y669_STAAM
55	39	45.9	295	1	AROE_CHLTE
56	39	45.9	327	1	K6P1_BACTN
57	39	45.9	339	1	SPSG_BACSU
58	39	45.9	382	1	GALI_SALTI
59	39	45.9	382	1	GALI_SALTY
60	39	45.9	512	1	GLGT_VICFA
61	39	45.9	678	1	GSPD_AERSA
62	39	45.9	679	1	UVRB_HAEIN
63	39	45.9	966	1	ENV_CAEVC
64	39	45.9	1173	1	GIGA_ORYSA
65	38	44.7	111	1	HV35_MOUSE
66	38	44.7	113	1	HV34_MOUSE
67	38	44.7	115	1	HV32_MOUSE
68	38	44.7	117	1	HV54_MOUSE
69	38	44.7	117	1	HV55_MOUSE
70	38	44.7	117	1	HV58_MOUSE
71	38	44.7	260	1	RS4_YARLI
72	38	44.7	261	1	RS4_DROME
73	38	44.7	262	1	RS4A_SCHPO
74	38	44.7	262	1	RS4B_SCHPO
75	38	44.7	262	1	RS4C_SCHPO
76	38	44.7	386	1	GALI_VIBPA
77	38	44.7	386	1	GALI_VIBVU
78	38	44.7	389	1	SERI_BOMMO
79	38	44.7	429	1	ASTI_YEAST
80	38	44.7	490	1	CAR3_DICDI
81	38	44.7	520	1	GLGS_BRANA
82	38	44.7	547	1	YQJN_BACSU
83	38	44.7	557	1	THS2_HALVO
84	38	44.7	562	1	HEMA_TAJAP
85	38	44.7	564	1	HEMA_TACKS
86	38	44.7	564	1	HEMA_TAMAP
87	38	44.7	565	1	HEMA_TATKO
88	38	44.7	609	1	HSF_SCHPO
89	38	44.7	942	1	AMPN_MANSE
90	38	44.7	1656	1	OMP8_PICJA
91	38	44.7	1666	1	MYM1_MOUSE
92	38	44.7	2127	1	RRPL_RABVS
93	38	44.7	2142	1	RRPL_RABVP
94	38	44.7	4289	1	TENX_HUMAN
95	37.5	44.1	97	1	HV56_MOUSE
96	37.5	44.1	206	1	URK_LACLA
97	37.5	44.1	248	1	PSPA_RAT
98	37.5	44.1	329	1	QOR_HUMAN
99	37.5	44.1	399	1	IRTF_MOUSE
100	37.5	44.1	718	1	SL53_BOVIN
101	37	43.5	121	1	FOLB_STAUA
102	37	43.5	139	1	NUSA_HALMO
103	37	43.5	146	1	HBB2_XENLA
104	37	43.5	146	1	HBB_PROHA
105	37	43.5	146	1	YQCK_BACSU
106	37	43.5	189	1	FAPR_LISIN

P47837	candida alb
Q8r003	mus musculus
Q9nu00	homo sapien
P30523	triticum ae
P55342	beta vulgar
P24116	vicia faba
P55238	hordeum vul
P55228	arabidopsis
P26140	influenza a
P52179	homo sapien
P19181	carassius a
Q01769	homo sapien
Q91755	streptococ
P49191	caenorhabdi
P15280	oryza sativ
Q64410	cavia porce
Q92119	ricetia
P26499	methanobact
P01812	mus musculus
Q8ctk9	staphylococ
Q8nkr2	staphylococ
Q8kbh8	chlorobium
Q8abrs	bacteroides
P39627	bacillus su
Q8z8b0	salmonella
P22713	salmonella
P52417	vicia faba
P45778	aeromonas s
P45125	haemophilus
P31626	caprine art
Q9awl7	oryza sativ
P01804	mus musculus
P01803	mus musculus
P01801	mus musculus
P18525	mus musculus
P18526	mus musculus
P18529	mus musculus
Q59950	yarrowia li
P41042	drosophila
P87158	schizosacch
Q9uaw5	schizosacch
Q9p4w9	schizosacch
Q87m60	vibrio para
Q8dbn9	vibrio vuln
P07856	bombyx mori
P35183	saccharomyc
P35352	dicyotateli
Q9m462	brassica na
P54551	bacillus su
Q30560	halobacteri
P03451	influenza a
P09345	influenza a
P87506	influenza a
P16060	influenza a
Q02953	schizosacch
P91885	manduca sex
Q06653	x outer mem
Q62234	mus musculus
P16289	rabies viru
P11213	rabies viru
P22105	homo sapien
P98521	mus musculus
Q9cf21	lactococcus
P08427	rattus norv
Q08257	homo sapien
P61179	mus musculus
P53793	bos taurus
P56740	staphylococ
P15738	halococcus
P02133	xenopus lae
P02086	procavia ca
P45945	bacillus su
Q92aj8	listeria in


```
QY      2 ISENGRTINYADSVKG 17
Db      51 ITWNGGSLYADSVKG 66

RESULT 3
HV3A_HUMAN
ID HV3A_HUMAN STANDARD; PRT; 122 AA.
AC P01762;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ig heavy chain V-III region TRO.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE (MYELOMA PROTEIN TRO).
RP MEDLINE=76023781; PubMed=809331;
RA Kratzin H., Altevogt P., Ruban E., Kortt A., Staroschik K.,
RA Hilschmann N.;
RT "The primary structure of a monoclonal IgA-immunoglobulin (IgA Tro.),
RT ii. The amino acid sequence of the H-chain, alpha-type, subgroup III;
RT structure of the complete IgA-molecule."
RL Hoppe-Seyler's Z. Physiol. Chem. 356:1337-1342(1975).
CC -!- MISCELLANEOUS: THE SEQUENCE OF THE C REGION IS ALSO GIVEN.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like domain.
DR PIR; A02045; ALHUTR.
DR HSSP; P01772; 2FB4.
DR GO; GO:0005576; C:extracellular; NAS.
DR GO; GO:0003823; F:antigen binding; NAS.
DR GO; GO:0006955; P:immune response; NAS.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; ig; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS0835; IG_LIKE; 1.
KW Immunoglobulin V region; Pyridolone carboxylic acid.
FT DOMAIN 1 108 IG-LIKE.
FT MOD_RES 1 108 PYRROLIDONE CARBOXYLIC ACID.
FT NON_TER 122 122
SQ SEQUENCE 122 AA; 13472 MW; 2E21A11DA04D80F9 CRC64;

Query Match 57.6%; Score 49; DB 1; Length 122;
Best Local Similarity 62.5%; Pred. No. 0.35;
Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY      2 ISENGRTINYADSVKG 17
Db      51 ICGSGSLYADSVKG 66

RESULT 4
HV16_MOUSE
ID HV16_MOUSE STANDARD; PRT; 136 AA.
AC P01763;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ig heavy chain V region MOPC 21 precursor (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RP MEDLINE=81234548; PubMed=6788376;
RA Bothwell A.L.M., Faskind M., Reith M., Imanishi-Kari T., Rajewsky K.,
RA Baltimore D.;
RT "Heavy chain variable region contribution to the NPb family of
RT antibodies: somatic mutation evident in a gamma 2a variable region.";

Cell 24:625-637(1981).
RN [2]
SEQUENCE OF 17-136.
RX MEDLINE=77100368; PubMed=401950;
RA Adetugbo K., Milstein C., Secher D.S.;
RT "Molecular analysis of spontaneous somatic mutants.";
RL Nature 265:299-304(1977).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J00522; AAD15290.1; -.
DR PIR; E90809; GIMS21.
DR PDB; 1IGC; 03-JUN-95.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; ig; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS0835; IG_LIKE; 1.
KW Immunoglobulin V region; Signal; 3D-structure.
FT NON_TER 1 16
FT SIGNAL <1 16
FT CHAIN 17 136 IG HEAVY CHAIN V REGION MOPC 21.
FT DOMAIN 115 119 D SEGMENT.
FT DOMAIN 120 136 JH4 SEGMENT.
FT DISULFID 38 112
FT CONFLICT 75 78 HYAD -> DYAH (IN REF. 2).
FT CONFLICT 89 90 DN -> ND (IN REF. 2).
FT CONFLICT 115 115 W -> H (IN REF. 2).
FT CONFLICT 120 120 Y -> W (IN REF. 2).
FT NON_TER 136 136
SQ SEQUENCE 136 AA; 15071 MW; 2276A98DBDF7016 CRC64;

Query Match 56.5%; Score 48; DB 1; Length 136;
Best Local Similarity 56.2%; Pred. No. 0.58;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY      2 ISENGRTINYADSVKG 17
Db      67 ISSGSLHYADTVKG 82

RESULT 5
HV3T_HUMAN
ID HV3T_HUMAN STANDARD; PRT; 116 AA.
AC P01781;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ig heavy chain V-III region GAL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE.
RX MEDLINE=75059123; PubMed=4803843;
RA Watanabe S., Barnikol H.U., Horn J., Bertram J., Hilschmann N.;
RT "The primary structure of a monoclonal IgM-immunoglobulin
RT (macroglobulin Gal.), II: the amino acid sequence of the H-chain (mu-
RT type), subgroup H III. Architecture of the complete IgM-molecule.";
RL Hoppe-Seyler's Z. Physiol. Chem. 354:1505-1509(1973).
RN [2]
REVISION TO 28-33.
RP Hilschmann N.;
RL Submitted (JUN-1975) to the PIR data bank.
CC -!- MISCELLANEOUS: THIS MU CHAIN WAS ISOLATED FROM A WALDENSTROM'S
CC MACROGLOBULIN.
```

```
CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
DR PIR; A02064; M3HUGL.
DR HSSP; P01772; 2FB4.
DR GO; GO:0005576; C:extracellular; NAS.
DR GO; GO:0003823; P:antigen binding; NAS.
DR GO; GO:0006955; P:immune response; NAS.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS0835; IG LIKE; 1.
KW Immunoglobulin V region.
FT DOMAIN 1 112 IG-LIKE.
FT NON_TER 116
SQ SEQUENCE 116 AA; 12730 MW; 2C67CA9AAAA1282 CRC64;

Query Match 55.3%; Score 47; DB 1; Length 116;
Best Local Similarity 50.0%; Pred. No. 0.72;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVK 17
DB 51 IKZBGSSZBYVDVSK 66

RESULT 6
HV03_CARAU STANDARD; PRT; 117 AA.
AC P19180;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V region 3 precursor.
OS Carassius auratus (Goldfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Carassius.
OX NCBI_TaxID=7957;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8814476; PubMed=3125551;
RA Wilson M.R., Middleton D., Warr G.W.;
RT "Immunoglobulin heavy chain variable region gene evolution: structure and family relationships of two genes and a pseudogene in a teleost fish."
RT Proc. Natl. Acad. Sci. U.S.A. 85:1566-1570(1988).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J03616; AA50807.1; -.
DR PIR; A28966; A28966.
DR HSSP; P01772; 2FB4.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS0835; IG LIKE; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 19
FT CHAIN 20 117 IG HEAVY CHAIN V REGION 3.
FT DOMAIN 20 49 FRAMEWORK-1.
FT DOMAIN 50 54 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 55 68 FRAMEWORK-2.
FT DOMAIN 69 85 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 86 117 FRAMEWORK-3.
FT DISULFID 41 115 BY SIMILARITY.
FT NON_TER 117
```

```
SQ SEQUENCE 117 AA; 13220 MW; 512B625003FA6ECB CRC64;

Query Match 54.1%; Score 46; DB 1; Length 117;
Best Local Similarity 50.0%; Pred. No. 1.1;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVK 17
DB 70 IYDDGSDVSADTVKG 85

RESULT 7
HV3G_HUMAN STANDARD; PRT; 122 AA.
ID HV3G_HUMAN
AC P01768;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ig heavy chain V-III region CAM.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=81013859; PubMed=6774332;
RA Lehman D.W., Putnam F.W.;
RT "Amino acid sequence of the variable region of a human mu chain: location of a possible JH segment."
RT Proc. Natl. Acad. Sci. U.S.A. 77:3239-3243(1980).
CC -1- MISCELLANEOUS: THIS MU CHAIN WAS ISOLATED FROM THE PLASMA OF A
CC PATIENT WITH MACROGLOBULINEMIA.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
DR PIR; A02051; M3HUM.
DR HSSP; P01772; 2FB4.
DR GO; GO:0005576; C:extracellular; NAS.
DR GO; GO:0003823; P:antigen binding; NAS.
DR GO; GO:0006955; P:immune response; NAS.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS0835; IG LIKE; 1.
KW Immunoglobulin V region; Pyrrolidone carboxylic acid.
FT DOMAIN 1 112 IG-LIKE.
FT MOD_RES 1 122 PYRROLIDONE CARBOXYLIC ACID.
FT NON_TER 122
SQ SEQUENCE 122 AA; 13668 MW; A42D0F17D252F1C2 CRC64;

Query Match 54.1%; Score 46; DB 1; Length 122;
Best Local Similarity 58.8%; Pred. No. 1.1;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 VISENGRTINYADSVK 17
DB 50 VISYGBBKXYABSVK 66

RESULT 8
HV3M_HUMAN STANDARD; PRT; 119 AA.
ID HV3M_HUMAN
AC P01774;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ig heavy chain V-III region POM.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=75046755; PubMed=4139708;
```

RA Capra J.D., Kehoe J.M.;
RT "Structure of antibodies with shared idiotype: the complete sequence
RT of the heavy chain variable regions of two immunoglobulin M
RT anti-gamma globulins.";
RL Proc. Natl. Acad. Sci. U.S.A. 71:4032-4036(1974).
CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM IGM WITH ANTI-GAMMA
CC GLOBULIN ACTIVITY.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
DR PIR; A02057; M3HUPM.
DR HSSP; P01772; 2FB4.
DR GO; GO:0005576; C:extracellular; NAS.
DR GO; GO:0003823; F:antigen binding; NAS.
DR GO; GO:0006955; P:immune response; NAS.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG-LIKE; 1.
KW Immunoglobulin V region.
FT DOMAIN 1 112
FT VARIANT 54 54 IG-LIKE.
FT N -> D (PROBABLY DUE TO DEAMIDATION
FT DURING ISOLATION).
FT /FTId=VAR_003966.
FT
FT
FT
FT
SQ NON_TER 119 119
SQ SEQUENCE 119 AA; 12953 MW; 2E018AF4DCB2610 CRC64;

Query Match 52.9%; Score 45; DB 1; Length 119;
Best Local Similarity 64.3%; Pred. No. 1.6;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 4 ENGRNTYADSVKG 17
DB 53 ENGNDRHYADSVNG 66
||| :||| |

RESULT 9
HV3N_HUMAN STANDARD; PRT; 119 AA.
ID HV3N_HUMAN
AC P01775;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ig heavy chain V-III region LAY.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=75046755; PubMed=4139708;
RA Capra J.D., Kehoe J.M.;
RT "Structure of antibodies with shared idiotype: the complete sequence
RT of the heavy chain variable regions of two immunoglobulin M
RT anti-gamma globulins.";
RL Proc. Natl. Acad. Sci. U.S.A. 71:4032-4036(1974).
CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM IGM WITH ANTI-GAMMA
CC GLOBULIN ACTIVITY.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
DR PIR; A02058; M3HULY.
DR HSSP; P01772; 2FB4.
DR GO; GO:0005576; C:extracellular; NAS.
DR GO; GO:0003823; F:antigen binding; NAS.
DR GO; GO:0006955; P:immune response; NAS.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG-LIKE; 1.
KW Immunoglobulin V region.
FT DOMAIN 1 112 IG-LIKE.
FT NON_TER 119 119
FT SEQUENCE 119 AA; 12858 MW; D6338098794DCF5E CRC64;

Query Match 52.9%; Score 45; DB 1; Length 119;
Best Local Similarity 64.3%; Pred. No. 1.6;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 4 ENGRNTYADSVKG 17
DB 53 ENGNDRHYADSVNG 66
||| :||| |

RESULT 10
HV3K_HUMAN STANDARD; PRT; 126 AA.
ID HV3K_HUMAN
AC P01772;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ig heavy chain V-III region KOL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE, AND DISULFIDE BONDS.
RX MEDLINE=83289131; PubMed=6884994;
RA Schmidt W.E., Jung H.-D., Palm W., Hilschmann N.;
RT "Three-dimensional structure determination of antibodies. Primary
RT structure of crystallized monoclonal immunoglobulin IgG1 KOL, I.";
RL Hoppe-Seyler's Z. Physiol. Chem. 364:713-747(1983).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX MEDLINE=81072295; PubMed=7441755;
RA Marquart M., Deisenhofer J., Huber R., Palm W.;
RT "Crystallographic refinement and atomic models of the intact
RT immunoglobulin molecule KOL and its antigen-binding fragment at 3.0 A
RT and 1.0-A resolution.";
RL J. Mol. Biol. 141:369-391(1980).
CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
DR PIR; A02055; GIHUKL.
DR PDB; 2FB4; 12-JUL-89.
DR PDB; 2IG2; 12-JUL-89.
DR GO; GO:0005576; C:extracellular; NAS.
DR GO; GO:0003823; F:antigen binding; NAS.
DR GO; GO:0006955; P:immune response; NAS.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG-LIKE; 1.
KW Immunoglobulin V region; 3D-structure; Pyrrolidone carboxylic acid.
FT DOMAIN 1 112 IG-LIKE.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT DISULFID 22 96
FT DISULFID 105 110
FT STRAND 3 7
FT STRAND 11 12
FT STRAND 14 15
FT STRAND 18 25
FT HELIX 29 31
FT STRAND 34 39
FT TURN 41 42
FT STRAND 45 51
FT STRAND 53 54
FT STRAND 58 60
FT HELIX 62 64
FT STRAND 65 65
FT TURN 66 67
FT STRAND 68 73
FT TURN 74 77
FT STRAND 78 83
FT HELIX 88 90
FT STRAND 92 99
FT STRAND 106 106
FT TURN 107 108

DR HSSP; P01810; 2FBJ.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00835; IG_LIKE; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 19
FT CHAIN 20 117 IG HEAVY CHAIN V REGION G4.
FT DOMAIN 20 49 FRAMEWORK-1.
FT DOMAIN 50 54 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 55 68 FRAMEWORK-2.
FT DOMAIN 69 85 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 86 117 FRAMEWORK-3.
FT DISULFID 41 115 BY SIMILARITY.
FT NON_TER 117 117
SQ SEQUENCE 117 AA; 12994 MW; 6330D7469AAE55FC CRC64;

Query Match 50.6%; Score 43; DB 1; Length 117;
Best Local Similarity 56.2%; Pred. No. 3.4;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVKG 17
| : | | | | |
Db 70 INGNSETIRYAPEVKG 85

RESULT 13
HV31_HUMAN
ID HV31_HUMAN STANDARD; PRT; 119 AA.
AC P01770;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ig heavy chain V-III region NIE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=77070269; PubMed=826475;
RA Ponatting H., Hilschmann N.;
RT "The rule of antibody structure. The primary structure of a
RT monoclonal IgG1 immunoglobulin (myeloma protein Nie). III. The
RT chymotryptic peptides of the H-chain, alignment of the tryptic
RT peptides and discussion of the complete structure.";
RL Hoppe-Seyler's Z. Physiol. Chem. 357:1571-1604(1976).
RN [2]
RP DISULFIDE BOND.
RX MEDLINE=77070267; PubMed=1002129;
RA Dreker L., Schwarz J., Reichel W., Hilschmann N.;
RT "Rule of antibody structure. The primary structure of a monoclonal
RT IgG1 immunoglobulin (myeloma protein Nie). I: Purification and
RT characterization of the protein, the L- and H-chains, the
RT cyanogen bromide cleavage products, and the disulfide bridges.";
RL Hoppe-Seyler's Z. Physiol. Chem. 357:1515-1540(1976).
CC -!- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IG1 MYELOMA
CC PROTEIN.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like domain.
CC PIR; A91668; GHUNI.
DR HSSP; P01772; 2FB4.
DR GO; GO:0005576; C:extracellular; NAS.
DR GO; GO:0003823; F:antigen binding; NAS.
DR GO; GO:0008955; F:immune response; NAS.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00835; IG_LIKE; 1.
KW Immunoglobulin V region; Pyroliadone carboxylic acid.
FT DOMAIN 1 112 IG-LIKE.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.

FT DISULFID 22 96
FT NON_TER 119 119
SQ SEQUENCE 119 AA; 13242 MW; C969355A6E55E165B CRC64;

Query Match 50.6%; Score 43; DB 1; Length 119;
Best Local Similarity 52.9%; Pred. No. 3.5;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 VISENGRTINYADSVKG 17
| : | | | | |
Db 50 VMSYGBGBKHADSVNG 66

RESULT 14
GALL_YERPE
ID GALL_YERPE STANDARD; PRT; 383 AA.
AC Q8ZG73;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Galactokinase (EC 2.7.1.6) (Galactose kinase).
GN GALK OR YP01137 OR Y3045.
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CO-92 / Biovar Orientalis;
RX MEDLINE=21470413; PubMed=11586360;
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Feltwell T., Hamlin N., Holroyd S., Jagers K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M.A., Rutherford K.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague.";
RL Nature 413:523-527(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=KIMS / Biovar Mediaevalis;
RX MEDLINE=2137863; PubMed=12142430;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA Perry R.D.;
RT "Genome sequence of Yersinia pestis KIM.";
RL J. Bacteriol. 184:4601-4611(2002).
CC -!- CATALYTIC ACTIVITY: ATP + D-galactose = ADP + D-galactose 1-
CC phosphate.
CC -!- PATHWAY: Galactose metabolism; first step.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -!- SIMILARITY: Belongs to the GHMP kinase family. Galk subfamily.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AJ414146; CAC89979.1; -;
CC DR EMBL; AE013905; AAC86596.1; -;
CC PIR; AH0139; AH0139.
CC HAMAP; MF 00246; -; 1.
CC InterPro; IPR000705; Galactokinase.
CC InterPro; IPR001174; Galkinase.
CC InterPro; IPR006204; GHMP kinase.
CC InterPro; IPR006203; GHMPKase ATP.
CC InterPro; IPR006206; Mgv_galkinase.

```
DR Pfam: PF00288; GHMP kinases; 1.
DR PRINTS; PR00473; GALTOKINASE.
DR PRINTS; PR00960; LMBPCKINASE.
DR PRINTS; PR00959; MEVGALKINASE.
DR TIGRFAMs; TIGR00131; gal kin; 1.
DR PROSITE; PS00106; GALTOKINASE; 1.
DR PROSITE; PS00627; GHMP KINASES ATP; 1.
KW Transferrase; Kinase; Galactose metabolism; ATP-binding;
KW Complete proteome.
FT NP_BIND 122 132 ATP (POTENTIAL).
SQ SEQUENCE 383 AA; 41865 MW; 548D19EB39D13281 CRC64;

Query Match 50.6%; Score 43; DB 1; Length 383;
Best Local Similarity 60.0%; Pred. No. 12;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 VISENARTLYADSV 15
DB 268 VISENARTLYADSV 282

RESULT 15
NADB NEIMA
ID NADB NEIMA STANDARD; PRT; 502 AA.
AC Q9JSX4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE L-aspartate oxidase (EC 1.4.3.16) (LASPO) (Quinolinate synthetase B).
GN NADB OR NMA2092.
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=22491 / Serogroup A / Serotype 4A;
RX MEDLINE=20222556; PubMed=10761919;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
RA Jagels K., Leacher S., Moule S., Mungall K., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrall B.G.;
RA "Complete DNA sequence of a serogroup A strain of Neisseria
RT meningitidis Z2491."
RL Nature 404:502-506(2000).
CC -1- FUNCTION: Catalyzes the oxidation of L-aspartate to
CC iminoaspartate.
CC -1- CATALYTIC ACTIVITY: L-aspartate + H(2)O + O(2) = oxaloacetate +
CC NH(3) + H(2)O(2).
CC -1- COFACTOR: FAD.
CC -1- PATHWAY: NAD biosynthesis; aspartate to NAMN; first step.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE FAD-DEPENDENT OXIDOREDUCTASE FAMILY 2.
CC NADB SUBFAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; AL162758; CAB85308.1; -.
CC PIR; E81780; E81780.
CC HSSP; P10902; 1CHU.
CC InterPro; IPR003953; FAD_bind2.
CC InterPro; IPR005288; NADB.
CC Pfam; PF00890; FAD_binding_2; 1.
CC TIGRFAMs; TIGR00551; nadB; 1.
KW Pyridine nucleotide biosynthesis; Oxidoreductase; Flavoprotein; FAD;

KW Pyridine nucleotide biosynthesis; Oxidoreductase; Flavoprotein; FAD;
KW Complete proteome.
FT NP_BIND 8 22 FAD (AMP PART) (POTENTIAL).
FT ACT_SITE 228 228 BY SIMILARITY.
FT ACT_SITE 249 249 BY SIMILARITY.
SQ SEQUENCE 502 AA; 54657 MW; 039EC390B823BC34 CRC64;

Query Match 50.6%; Score 43; DB 1; Length 502;
Best Local Similarity 53.3%; Pred. No. 16;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 3 SENGRTINYADSVK 17
DB 236 SENGRTINYADSVK 250

RESULT 16
NADB NEIMB
ID NADB NEIMB STANDARD; PRT; 502 AA.
AC Q9KJ07;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE L-aspartate oxidase (EC 1.4.3.16) (LASPO) (Quinolinate synthetase B).
GN NADB OR NMB0392.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MC58 / Serogroup B;
RX MEDLINE=20175755; PubMed=10710107;
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Ciecko A., Parksey D.S., Blair E., Citron H., Vanthaeven J.,
RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
RA Gill J., Scarlato V., Masingani V., Pizzo M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RA "Complete genome sequence of Neisseria meningitidis serogroup B strain
RT MC58."
RL Science 287:1809-1815(2000).
CC -1- FUNCTION: Catalyzes the oxidation of L-aspartate to
CC iminoaspartate.
CC -1- CATALYTIC ACTIVITY: L-aspartate + H(2)O + O(2) = oxaloacetate +
CC NH(3) + H(2)O(2).
CC -1- COFACTOR: FAD.
CC -1- PATHWAY: NAD biosynthesis; aspartate to NAMN; first step.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE FAD-DEPENDENT OXIDOREDUCTASE FAMILY 2.
CC NADB SUBFAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; AE002395; AAP40832.1; -.
CC PIR; H81203; H81203.
CC HSSP; P10902; 1CHU.
CC TIGR; NME0392; -.
CC InterPro; IPR003953; FAD_bind2.
CC InterPro; IPR005288; NADB.
CC InterPro; IPR004112; Succ_DH_flav_C.
CC Pfam; PF00890; FAD_binding_2; 1.
CC TIGRFAMs; TIGR00551; nadB; 1.
KW Pyridine nucleotide biosynthesis; Oxidoreductase; Flavoprotein; FAD;
KW Complete proteome.
```

```

FT NP BIND      8      22      FAD (AMP PART) (POTENTIAL).
FT ACT SITE    228     228      BY SIMILARITY.
FT ACT SITE    249     249      BY SIMILARITY.
SQ SEQUENCE    502 AA; 54562 MW; 644152CC3753E294 CRC64;

Query Match      50.6%; Score 43; DB 1; Length 502;
Best Local Similarity 53.3%; Pred. No. 16;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 3 SENGRTINVDVSKG 17
||||| :||:|
Db 236 SENGRTFLISEAVRG 250

RESULT 17
UVRB LISIN STANDARD; PRT; 658 AA.
AC Q28A4;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE UvrABC system protein B (UvrB protein) (Excinuclease ABC subunit B).
GN UVRB OR LIN2632.
OS Listeria innocua.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1642;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIP 11262 / Serovar 6a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaeser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fshih H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kretz J., Kuhn M., Kunst F., Kurapkat G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
CC -!- FUNCTION: The UvrABC repair system catalyzes the recognition and
CC processing of DNA lesions. A damage recognition complex composed
CC of 2 uvrA and 2 uvrB subunits scans DNA for abnormalities. Upon
CC binding of the uvrA(2)B(2) complex to a putative damaged site, the
CC DNA wraps around one uvrB monomer. DNA wrap is dependent on ATP
CC binding by uvrB and probably causes local melting of the DNA
CC helix, facilitating insertion of uvrB beta-hairpin between the DNA
CC strands. Then uvrB probes one DNA strand for the presence of a
CC lesion. If a lesion is found the uvrA subunits dissociate and the
CC uvrB-DNA preincision complex is formed. This complex is
CC subsequently bound by uvrC and the second uvrB is released. If no
CC lesion is found, the DNA wraps around the other uvrB subunit that
CC will check the other strand for damage (By similarity).
CC -!- SUBUNIT: Forms a heterotetramer with uvrA during the search for
CC lesions. Interacts with uvrC in an incision complex (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- DOMAIN: The beta-hairpin motif is involved in DNA binding (By
CC similarity).
CC -!- SIMILARITY: Belongs to the uvrB family.
CC -!- SIMILARITY: Contains 1 UVR domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial/
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

DR EMBL; AL596173; CAC97859.1; -.
DR PIR; AC1761; AC1761.
DR Listlist; LINO2632; -.
DR HAMAP; MF_00204; -.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004807; UvrB.
DR InterPro; IPR001943; UvrB/C.
DR Pfam; PF0271; helicase_C; 1.
DR Pfam; PF02151; Uvr; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELICC; 1.
DR TIGRFAMs; TIGR00631; uvrB; 1.
DR PROSITE; PS0151; UVR; 1.
KW SOS response; Excision nuclease; DNA repair; DNA recombination;
KW DNA excision; ATP-binding; Complete proteome.
FT NP_BIND 39 46 ATP (POTENTIAL).
FT DOMAIN 92 115 BETA-HAIRPIN.
FT DOMAIN 622 657 UVR.
SQ SEQUENCE 658 AA; 75616 MW; 7F15B338B0E3BB6F CRC64;

Query Match      50.6%; Score 43; DB 1; Length 658;
Best Local Similarity 61.5%; Pred. No. 22;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 SENGRTINVDVSKG 15
||||| :||:|
Db 544 NENGRVINYADKM 556

RESULT 18
UVRB LISMO STANDARD; PRT; 658 AA.
AC Q8YAF5;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE UvrABC system protein B (UvrB protein) (Excinuclease ABC subunit B).
GN UVRB OR LMO2469.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EGD-e / Serovar 1/2a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaeser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fshih H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kretz J., Kuhn M., Kunst F., Kurapkat G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
CC -!- FUNCTION: The UvrABC repair system catalyzes the recognition and
CC processing of DNA lesions. A damage recognition complex composed
CC of 2 uvrA and 2 uvrB subunits scans DNA for abnormalities. Upon
CC binding of the uvrA(2)B(2) complex to a putative damaged site, the
CC DNA wraps around one uvrB monomer. DNA wrap is dependent on ATP
CC binding by uvrB and probably causes local melting of the DNA
CC helix, facilitating insertion of uvrB beta-hairpin between the DNA
CC strands. Then uvrB probes one DNA strand for the presence of a
CC lesion. If a lesion is found the uvrA subunits dissociate and the
CC uvrB-DNA preincision complex is formed. This complex is
CC subsequently bound by uvrC and the second uvrB is released. If no
CC lesion is found, the DNA wraps around the other uvrB subunit that
CC will check the other strand for damage (By similarity).
CC -!- SUBUNIT: Forms a heterotetramer with uvrA during the search for
CC lesions. Interacts with uvrC in an incision complex (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- DOMAIN: The beta-hairpin motif is involved in DNA binding (By
CC similarity).
CC -!- SIMILARITY: Belongs to the uvrB family.
CC -!- SIMILARITY: Contains 1 UVR domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial/
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

CC lesions. Interacts with uvrC in an incision complex (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- DOMAIN: The beta-hairpin motif is involved in DNA binding (By
CC similarity).
CC -!- SIMILARITY: Belongs to the uvrB family.
CC -!- SIMILARITY: Contains 1 UVR domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AL591983; CAD00567.1; -.
CC PIR; A11385; A11385
CC ListList; LMO02489; -.
CC HAMAP; MF_00204; -.
CC InterPro; IPR001410; DEAD.
CC InterPro; IPR001650; Helicase_C.
CC InterPro; IPR004807; UvrB.
CC InterPro; IPR001943; UvrB/C.
CC Pfam; PF00271; helicase_C; 1.
CC Pfam; PF02151; UVR; 1.
CC SMART; SM00487; DEXDC; 1.
CC SMART; SM00490; HELIC; 1.
CC TIGRFAMs; TIGR00631; uvrB; 1.
CC PROSITE; PS50151; UVR; 1.
CC SOS response; Excision nuclease; DNA repair; DNA recombination;
KW DNA excision; ATP-binding; Complete proteome.
FT NP BIND 39 46
FT DOMAIN 92 115
FT DOMAIN 622 657
FT SEQUENCE 658 AA; 75562 MW; AA257263C2C71919 CRC64;
SQ
Query Match 50.6%; Score 43; DB 1; Length 658;
Best Local Similarity 61.5%; Pred. No. 22;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 SENGRTINVADSV 15
Db 544 NENGRVIMYADM 556
:|||||:
:|||||:

RESULT 19
PHYB SOYB
ID PHYB_SOYB STANDARD; PRT; 1156 AA.
AC P42459;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Phytochrome B.
GN PHYB.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=cv. Paldal; TISSUE=Etioolated leaf;
RC Hahn T.R., Woo T.W., Seo H.S., Choi Y.D.;
RL Submitted (XXX-1994) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Regulatory photoreceptor which exists in two forms that
CC are reversibly interconvertible by light: the Pr form that absorbs
CC maximally in the red region of the spectrum and the Pfr form that
CC absorbs maximally in the far-red region. Photoconversion of Pr in
CC Pfr induces an array of morphogenic responses, whereas
CC reconversion of Pfr to Pr cancels the induction of those
CC responses. Pfr controls the expression of a number of nuclear
CC genes including those encoding the small subunit of ribulose-

```

```

CC bisphosphate carboxylase, chlorophyll A/B binding protein,
CC protochlorophyllide reductase, rRNA, etc. It also controls the
CC expression of its own gene(s) in a negative feedback fashion.
CC -!- SUBUNIT: Homodimer.
CC -!- PTM: Contains one covalently linked tetrapyrrole chromophore.
CC -!- SIMILARITY: Belongs to the phytochrome family.
CC -!- SIMILARITY: Contains 1 PAS (PER-ARNT-SIM) dimerization domain.
CC -!- SIMILARITY: Contains 1 histidine kinase domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L34843; AAA34000.1; -.
CC PIR; T07756; T07756.
CC InterPro; IPR003594; ATPbind_ATPase.
CC InterPro; IPR003018; GAF.
CC InterPro; IPR003661; His_kinA_N.
CC InterPro; IPR005467; His_kinase.
CC InterPro; IPR000014; PAS_domain.
CC InterPro; IPR001294; Phytochrome.
CC Pfam; PF01590; GAF; 1.
CC Pfam; PF02518; HATPase_C; 1.
CC Pfam; PF00512; HSKA; 1.
CC Pfam; PF00989; PAS; 2.
CC Pfam; PF00360; phytochrome; 1.
CC PRINTS; PR01033; PHYTOCHROME.
CC SMART; SM00065; GAF; 1.
CC SMART; SM00387; HATPase_C; 1.
CC SMART; SM00388; HSKA; 1.
CC SMART; SM00091; PAS; 1.
CC TIGRFAMs; TIGR00229; sensory_box; 1.
CC PROSITE; PS50109; HIS_KIN; 1.
CC PROSITE; PS50112; PAS; 1.
CC PROSITE; PS00245; PHYTOCHROME_1; 1.
CC PROSITE; PS00446; PHYTOCHROME_2; 1.
CC Transcription regulation; Photoreceptor; Phytochrome; Chromophore;
KW Multigene family.
FT DOMAIN 646 717
FT DOMAIN 930 1150
FT BINDING 358 358
FT SEQUENCE 1156 AA; 129085 MW; 76333AABDC42D297 CRC64;
SQ
Query Match 50.6%; Score 43; DB 1; Length 1156;
Best Local Similarity 46.2%; Pred. No. 40;
Matches 6; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

QY 4 ENGRNTINVADSVK 16
Db 59 ESGRSFNYSIESIR 71
:|||||:
:|||||:

RESULT 20
HV02 CANFA
ID HV02 CANFA STANDARD; PRT; 117 AA.
AC P01785;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ig heavy chain V region MOO.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
SEQUENCE OF 1-112.
RP MEDLINE=77242268; PubMed=407924;
RX Wasserman R.L., Capra J.D.;
RA "Primary structure of the variable regions of two canine
RT

```

immunoglobulin heavy chains.";
 RN Biochemistry 16:3160-3168(1977).
 [2]
 RP SEQUENCE OF 113-117.
 RX MEDLINE=80077682; PubMed=1117299;
 RA McCumber L.J., Capra J.D.;
 RT "The complete amino-acid sequence of a canine mu chain.";
 RL Mol. Immunol. 16:565-570(1979).
 CC -!- MISCELLANEOUS: THIS MU CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like domain.
 DR PIR; A90403; MHDGMO.
 DR HSSP; P01772; 2FB4.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003596; Ig_v.
 DR Pfam; PF00047; Ig; 1.
 DR SMART; SM00406; IGV; 1.
 DR PROSITE; PS50835; IG LIKE; 1.
 KW Immunoglobulin V region.
 FT DOMAIN 1 116 IG-LIKE.
 FT NON_TER 117 117
 SQ SEQUENCE 117 AA; 12703 MW; FCE07309E0A84B35 CRC64;
 Query Match 50.0%; Score 42.5; DB 1; Length 117;
 Best Local Similarity 62.5%; Pred. No. 4.1;
 Matches 10; Conservative 3; Mismatches 2; Indels 1; Gaps 1;
 QY 2 ISENGRTINYADSVKG 17
 DB 51 ISSSGQTY-YADAVKG 65
 RESULT 21
 ID HV3F HUMAN STANDARD; PRT; 115 AA.
 AC P01767;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Ig heavy chain V-III region BUT.
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=78137069; PubMed=416441;
 RA Torano A., Putnam F.W.;
 RT "Complete amino acid sequence of the alpha 2 heavy chain of a human
 RL IgA2 immunoglobulin of the A2m (2) allotype.";
 RL Proc. Natl. Acad. Sci. U.S.A. 75:966-969(1978).
 CC -!- MISCELLANEOUS: THE SEQUENCE OF THE ALPHA-2, A2M(2) ALLOTYPE, C
 CC REGION OF THIS MYELOMA PROTEIN IS ALSO GIVEN.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like domain.
 DR PIR; A02050; AZHUBU.
 DR HSSP; P01789; 1MCP.
 DR GO; GO:0005576; C:extracellular; NAS.
 DR GO; GO:0003823; F:antigen binding; NAS.
 DR GO; GO:0006955; P:immune response; NAS.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003596; Ig_v.
 DR Pfam; PF00047; Ig; 1.
 DR SMART; SM00406; IGV; 1.
 DR PROSITE; PS50835; IG LIKE; 1.
 KW Immunoglobulin V region.
 FT DOMAIN 1 111 IG-LIKE.
 FT NON_TER 115 115
 SQ SEQUENCE 115 AA; 12379 MW; 208876A7DF52DCF4 CRC64;
 Query Match 49.4%; Score 42; DB 1; Length 115;
 Best Local Similarity 75.0%; Pred. No. 4.9;
 Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 6 GRTINYADSVKG 17

Db 54 GGTYYADSVKG 65
 RESULT 22
 ID HV3J HUMAN STANDARD; PRT; 121 AA.
 AC P01771;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Ig heavy chain V-III region HIL.
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=79124695; PubMed=420800;
 RA Chiu Y.-Y.H., Lopez de Castro J.A., Poljak R.J.;
 RT "Amino acid sequence of the VH region of human myeloma
 RT cryoimmunoglobulin IgG Hil.";
 RL Biochemistry 18:553-560(1979).
 CC -!- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGG1 MYELOMA
 CC PROTEIN.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like domain.
 DR PIR; A02054; GIHULH.
 DR HSSP; P01772; 2FB4.
 DR GO; GO:0005576; C:extracellular; NAS.
 DR GO; GO:0003823; F:antigen binding; NAS.
 DR GO; GO:0006955; P:immune response; NAS.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003596; Ig_v.
 DR Pfam; PF00047; IGV; 1.
 DR SMART; SM00406; IGV; 1.
 DR PROSITE; PS50835; IG LIKE; 1.
 KW Immunoglobulin V region; Pyrrolidone carboxylic acid.
 FT DOMAIN 1 112 IG-LIKE.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT NON_TER 121 121
 SQ SEQUENCE 121 AA; 13566 MW; 480FCS3610EFD5DAB CRC64;
 Query Match 49.4%; Score 42; DB 1; Length 121;
 Best Local Similarity 58.8%; Pred. No. 5.2;
 Matches 10; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 1 VISENGRTINYADSVKG 17
 DB 50 VIWYNGSRITYGDSVKG 66
 RESULT 23
 ID HAMI BORBU STANDARD; PRT; 201 AA.
 AC O51263;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE HAMI protein homolog.
 GN BB0247.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxID=139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 35210 / B31;
 RX MEDLINE=98065943; PubMed=9403685;
 RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
 RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
 RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
 RA Peterson J., Kerkavag A.R., Quackenbush J., Salzberg S., Hanson M.,
 RA van Vugt R., Palmer N., Adams M.D., Cocayne J.D., Weidman J.,
 RA Utterback T., Wattley L., McDonald L., Attiach P., Bowman C.,

```

RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
RT burgdorferi.";
RL Nature 390:580-586(1997).
CC -1- SIMILARITY: Belongs to the HAM1 NTPase family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AE001135; AAB91499.1; -.
DR F1R; G70130; G70130.
DR HSP; Q57679; I878.
DR TIGR; BB0247; -.
DR HAMAP; MF 01405; -. 1.
DR InterPro; IPR002637; Hampl like.
DR Ffam; PF01725; Hampl like_1
DR TIGRfams; TIGR00042; TIGR00042; 1.
DR Hydrolase; Complete proteome.
KW SEQUENCE 201 AA; 23138 MW; 199E43E0704B4575 CRC64;
-----
Query Match 49.4%; Score 42; DB 1; Length 201;
Best Local Similarity 43.8%; Pred. No. 9;
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
-----
QY 2 ISENGRTINYADSVKG 17
DB 126 ISKNGQILNFEIGIK 141
|||:::|:|:|
-----
RESULT 24
MBNL MOUSE
ID ID_MBNL_MOUSE STANDARD; PRT; 341 AA.
AC Q9JK5;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Muscblind-like protein (Triplet-expansion RNA-binding protein).
DE MBNL1 OR MBNL OR EXP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RA Castagnola P., Monticone M., Borsani G., Bassi M.T., Tonachini L.;
RT "cDNA cloning of mouse muscblind.",
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Binds to CUG triplet repeat expansion dsRNA (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- SIMILARITY: Contains 4 C3H1-type zinc fingers.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AF231110; AAF72159.1; -.
DR MGD; MGI.1928482; Mbn1l.
DR GO; GO:0005737; Cytoplasm; ISS.
DR GO; GO:0005634; C:nucleus; ISS.
DR GO; GO:0003725; F:double-stranded RNA binding; ISS.
DR GO; GO:0001701; P:embryonic development (sensu Mammalia); IEP.
DR GO; GO:000326; P:limb morphogenesis; IEP.

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; Y13829; CAAY74155.1; -;
 DR EMBL; AB007888; BAA24858.1; -;
 DR EMBL; AF255334; AAF76138.1; -;
 DR Genew; HGNC:6923; MBNL1.
 DR MIM; 606516; -;
 DR GO; GO:0005737; C:cytoplasm; IDA.
 DR GO; GO:0005634; C:nucleus; IDA.
 DR GO; GO:0003725; F:double-stranded RNA binding; IDA.
 DR GO; GO:0001701; P:embryonic development (sensu Mammalia); ISS.
 DR GO; GO:0030326; P:limb morphogenesis; ISS.
 DR GO; GO:0007517; P:muscle development; ISS.
 DR GO; GO:0045445; P:myoblast differentiation; ISS.
 DR GO; GO:0007399; P:neurogenesis; ISS.
 DR InterPro; IPR000571; Znf.CCCH.
 DR Pfam; PF00642; zf-CCCH; 4.
 DR SMART; SM00356; Znf_C3H1; 4.
 DR Zinc-finger; Repeat; Nuclear protein; RNA-binding;
 KW Alternative splicing.
 FT ZN_FING 14 40 C3H1-TYPE 1.
 FT ZN_FING 47 72 C3H1-TYPE 2.
 FT ZN_FING 179 206 C3H1-TYPE 3.
 FT ZN_FING 216 240 C3H1-TYPE 4.
 FT VARSPPLIC 116 183 Missing (in isoform EXP35).
 FT VARSPPLIC 270 287 Missing (in isoform EXP35 and isoform
 FT Missing (in isoform EXP40).
 FT /FTId-VSP 006430.
 FT /FTId-VSP 006430.
 SQ SEQUENCE 388 AA; 41817 MW; 118D256A81A86695 CRC64;

Query Match 49.4%; Score 42; DB 1; Length 388;
 Best Local Similarity 64.3%; Pred. No. 18;
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 4 ENGRITNYADSVKKG 17
 DB 46 ENGRVIACFDLSLKG 59

Search completed: September 24, 2004, 01:55:23
 Job time : 16.1132 secs

This page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 01:47:17 ; Search time 81.4717 seconds
(without alignments)
65.836 Million cell updates/sec

Title: US-10-088-639a-2_COPY_177_193
Perfect score: 85
Sequence: 1 VISENGRTINADSVKG 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database : SPTREMBL_25.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Description
1	58	68.2	Q9APP6 homo sapien
2	54	63.5	Q8NSK4
3	50	58.8	Q9UL87
4	50	58.8	Q8WU38
5	48	56.5	Q9HCC1
6	47	55.3	Q9UL93
7	47	55.3	Q8KW71
8	47	55.3	Q8WU71
9	46	54.1	Q9UL71
10	46	54.1	Q8A8H8
11	46	54.1	Q88FX8
12	45	52.9	Q26670
13	45	52.9	Q96BB9
14	45	52.9	Q9BLH9
15	45	52.9	Q88LF9
16	45	52.9	Q8YC77

17	45	52.9	729	16	Q8FW49
18	45	52.9	853	15	Q9IV97
19	45	52.9	1105	16	Q9A430
20	45	52.9	3449	10	Q9ZQX5
21	44	51.8	118	4	Q9UL91
22	44	51.8	122	4	Q9UL84
23	44	51.8	220	14	Q9JH96
24	44	51.8	473	11	Q91Z05
25	44	51.8	624	16	Q7VJ23
26	44	51.8	504	10	Q48662
27	44	51.8	658	16	Q8IX47
28	44	51.8	658	16	Q8IS13
29	43.5	51.2	465	10	Q9FWW9
30	43	50.6	112	4	Q9UGP3
31	43	50.6	153	5	O45948
32	43	50.6	313	16	Q8P9V2
33	43	50.6	418	16	Q89AC6
34	43	50.6	576	10	Q8H2F6
35	43	50.6	2375	5	Q8IEN3
36	43	50.6	2391	5	Q27732
37	42.5	50.0	386	5	Q8SZ70
38	42.5	50.0	407	5	Q9VW14
39	42.5	50.0	548	16	Q9A4M8
40	42	49.4	114	3	Q9HEN3
41	42	49.4	147	4	Q9Y509
42	42	49.4	255	4	O95205
43	42	49.4	314	4	Q96RE3
44	42	49.4	339	11	Q8CAQ1
45	42	49.4	341	4	Q86UV9
46	42	49.4	342	4	Q86UV8
47	42	49.4	355	11	Q8C181
48	42	49.4	361	4	Q8NEV3
49	42	49.4	367	4	Q8TD82
50	42	49.4	373	11	Q8BUD8
51	42	49.4	382	4	Q96P92
52	42	49.4	384	11	Q8CHP5
53	42	49.4	421	2	Q7X2A8
54	42	49.4	482	4	Q7Z351
55	42	49.4	514	10	Q9LLU6
56	42	49.4	562	12	Q67185
57	42	49.4	671	16	Q884C9
58	42	49.4	741	16	Q8UFY1
59	42	49.4	921	16	Q8ECE9
60	42	49.4	1036	10	Q7XST9
61	42	49.4	1147	2	P94448
62	42	49.4	1154	4	Q9UPP6
63	42	49.4	1192	4	Q9UL68
64	42	49.4	1541	17	Q8TTP3
65	42	49.4	9271	5	Q8IES7
66	41	48.2	165	10	O9AQ27
67	41	48.2	293	16	Q8ZL44
68	41	48.2	294	16	Q8Z2H7
69	41	48.2	294	16	Q83SU0
70	41	48.2	444	10	Q43152
71	41	48.2	473	10	O9M4Z1
72	41	48.2	474	10	Q84PF5
73	41	48.2	474	10	Q84PF4
74	41	48.2	474	10	Q84PF3
75	41	48.2	474	10	Q84JDB
76	41	48.2	474	10	Q84J86
77	41	48.2	474	10	Q84J79
78	41	48.2	475	10	O947C0
79	41	48.2	480	11	Q91XE1
80	41	48.2	505	10	Q9AT05
81	41	48.2	507	10	Q43816
82	41	48.2	515	10	Q84UT1
83	41	48.2	516	10	O9AT06
84	41	48.2	517	10	O947B9
85	41	48.2	518	10	P93229
86	41	48.2	522	10	P93476
87	41	48.2	523	10	O9M4W7
88	41	48.2	523	10	P93477
89	41	48.2	525	10	O22629

Q8FW49	brucella su
Q9IV97	human immun
Q9A430	caulobacter
Q9ZQX5	arabidopsis
Q9UL91	homo sapien
Q9UL84	homo sapien
Q9JH96	unclassified
Q91Z05	mus musculus
O7VJ23	helicobacte
O48662	cucurbita c
Q8IX47	bacillus an
Q8IS13	bacillus ce
Q9FWW9	arabidopsis
O9UGP3	homo sapien
O45948	caenorhabdi
Q8P9V2	xanthomonas
Q89AC6	buchnera ap
Q8H2F6	oryza sativ
Q8IEN3	plasmodium
Q27732	plasmodium
Q8SZ70	drosophila
Q9VW14	drosophila
Q9A4M8	caulobacter
Q9HEN3	cochliobolu
O9Y509	homo sapien
O95205	homo sapien
O96RE3	homo sapien
Q8CAQ1	mus musculus
Q86UV9	homo sapien
Q86UV8	homo sapien
Q8C181	mus musculus
Q8NEV3	homo sapien
Q8TD82	homo sapien
Q8BUD8	mus musculus
Q96P92	homo sapien
Q8CHF5	mus musculus
Q7X2A8	gemmata sp.
Q7Z351	homo sapien
Q9LLU6	chlamydomon
Q67185	influenza a
Q884C9	pseudomonas
Q8UFY1	agrobacteri
Q8ECE9	shewanella
O7XST9	oryza sativ
P94448	bacillus st
Q9UPP6	homo sapien
Q9UL68	homo sapien
Q8TTP3	methanosarc
Q8IES7	plasmodium
O9AQ27	amorphophal
Q8ZL44	salmonella
Q8Z2H7	salmonella
Q83SU0	salmonella
Q43152	spinacia ol
O9M4Z1	trititum ae
Q84PF5	zea mays su
Q84PF4	zea mays su
Q84PF3	zea mays su
Q84JDB	zea mays su
Q84J86	zea mays su
Q84J79	zea mays su
O947C0	zea mays (m
Q91XE1	mus musculus
Q9AT05	cicer ariet
Q43816	pisum sativ
Q84UT1	phaseolus v
Q9AT06	cicer ariet
O947B9	zea mays (m
P93229	lycopersico
P93476	ipomoea bat
Q9M4W7	perilla fru
P93477	ipomoea bat
O22629	cucumis mel

```
90 41 48.2 526 10 022657 022657 citrullus 1
91 41 48.2 531 12 009642 009642 inflenzavi
92 41 48.2 541 12 092NC7 Q9ZNC7 clostridium
93 41 48.2 541 16 08XK6 Q8XK6 clostridium
94 41 48.2 565 12 08QT92 08QT92 inflenza a
95 41 48.2 565 12 091GSO 091GSO inflenza a
96 41 48.2 566 12 08QT94 08QT94 inflenza a
97 41 48.2 566 12 08UYCO 08UYCO inflenza a
98 41 48.2 566 12 08QI33 08QI33 inflenza a
99 41 48.2 566 12 08QV02 08QV02 inflenza a
100 41 48.2 566 12 08QT93 08QT93 inflenza a
101 41 48.2 566 12 08QT98 08QT98 inflenza a
102 41 48.2 566 12 08QT91 08QT91 inflenza a
103 41 48.2 566 12 08QI36 08QI36 inflenza a
104 41 48.2 566 12 08QV30 08QV30 inflenza a
105 41 48.2 566 12 08QT90 08QT90 inflenza a
106 41 48.2 566 12 091BX7 091BX7 inflenza a
107 41 48.2 566 12 08QPY3 08QPY3 inflenza a
108 41 48.2 566 12 091I38 091I38 inflenza a
109 41 48.2 566 12 08QI34 08QI34 inflenza a
110 41 48.2 566 12 08QT87 08QT87 inflenza a
111 41 48.2 566 12 08QV32 08QV32 inflenza a
112 41 48.2 566 12 08QV02 08QV02 inflenza a
113 41 48.2 566 12 091BX8 091BX8 inflenza a
114 41 48.2 566 12 08QI64 08QI64 inflenza a
115 41 48.2 566 12 08QI37 08QI37 inflenza a
116 41 48.2 566 12 08Q459 08Q459 inflenza a
117 41 48.2 566 12 08B2T6 08B2T6 inflenza a
118 41 48.2 566 12 08Q9F7 08Q9F7 inflenza a
119 41 48.2 593 16 08F470 08F470 leptospira
120 41 48.2 662 9 09G064 09G064 burkholderi
121 41 48.2 760 16 07V1Y7 07V1Y7 prochloroco
122 41 48.2 792 16 08RKS3 08RKS3 streptomyce
123 41 48.2 900 10 094B77 094B77 arabidopsis
124 41 48.2 900 10 09FIU0 09FIU0 arabidopsis
125 41 48.2 3013 16 08ZA73 08ZA73 yersinia pe
126 41 47.1 90 10 094JM6 094JM6 oryza sativ
127 40 47.1 118 4 09UL72 09UL72 homo sapien
128 40 47.1 135 2 08RJ70 08RJ70 streptococc
129 40 47.1 135 2 09R7Y7 09R7Y7 streptococc
130 40 47.1 135 2 09R7Y6 09R7Y6 streptococc
131 40 47.1 135 2 09R7Y5 09R7Y5 streptococc
132 40 47.1 135 2 09ZEF1 09ZEF1 streptococc
133 40 47.1 135 2 08R749 08R749 streptococc
134 40 47.1 135 2 09R7Y4 09R7Y4 streptococc
135 40 47.1 135 2 09R321 09R321 streptococc
136 40 47.1 135 2 09RCP6 09RCP6 streptococc
137 40 47.1 135 2 09ZEF0 09ZEF0 streptococc
138 40 47.1 135 2 09R394 09R394 streptococc
139 40 47.1 135 2 09R3E8 09R3E8 streptococc
140 40 47.1 135 2 09RCP5 09RCP5 streptococc
141 40 47.1 135 2 09R370 09R370 streptococc
142 40 47.1 135 2 09R7Y8 09R7Y8 streptococc
143 40 47.1 149 2 07X5G9 07X5G9 streptococc
144 40 47.1 149 2 07X5G7 07X5G7 streptococc
145 40 47.1 149 2 07X5G6 07X5G6 streptococc
146 40 47.1 149 2 07X5G5 07X5G5 streptococc
147 40 47.1 149 2 07X5G1 07X5G1 streptococc
148 40 47.1 149 2 07X5F7 07X5F7 streptococc
149 40 47.1 149 2 07X5F5 07X5F5 streptococc
150 40 47.1 149 2 07X5F4 07X5F4 streptococc
```

ALIGNMENTS

```
RESULT 1
Q9NPP6 PRELIMINARY; PRT; 416 AA.
ID Q9NPP6
AC Q9NPP6 (T-EMBLrel. 15, Created)
DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)

DE Immunoglobulin heavy chain variant (Fragment).
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Pluvinet R., Estivill X., Escarceller M., Sumoy L.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Auffray C., Ansoerge W., Ballabio A., Estivill X., Gibson K.,
RA Lehrach H., Poustka A., Lundeberg J.;
RT "The European IMAGE consortium for integrated Molecular analysis of
human gene transcripts.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL389978; CAB97534.1; -.
DR HSSP; P01789; 1MCP.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003006; IG_MHC.
DR Pfam; PF00047; IG; 4.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG LIKE; 3.
DR PROSITE; PS00290; IG_MHC; 2.
FT NON_TER
SQ SEQUENCE 416 AA; 44786 MW; 8C41708BB8AB4687 CRC64;

Query Match 68.2%; Score 58; DB 4; Length 416;
Best Local Similarity 68.8%; Pred. No. 0.34;
Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 2 ISENGRTINYADSVKG 17
||:|:|:|:|:|:|
Db 11 ISSSGDTVDYADSVKG 26

RESULT 2
Q8NSK4 PRELIMINARY; PRT; 499 AA.
ID Q8NSK4
AC Q8NSK4;
DT 01-OCT-2002 (T-EMBLrel. 22, Created)
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC032249; AAH32249.1; -.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003597; IG-cl.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003596; IG_v.
DR Pfam; PF00047; ig; 4.
DR SMART; SM00409; IG; 4.
DR SMART; SM00407; IGcl; 2.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG LIKE; 4.
DR PROSITE; PS00290; IG_MHC; 1.
KW Hypothetical protein.
SQ SEQUENCE 499 AA; 53376 MW; 93A5C89582054F32 CRC64;

Query Match 63.5%; Score 54; DB 4; Length 499;
Best Local Similarity 68.8%; Pred. No. 2;
Matches 11; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
```

Qy 2 ISENGRTINYADSVKG 17
 Db 70 INWNGSTGYADSVKG 85

RESULT 3

Q9UL87 Q9UL87 PRELIMINARY; PRT; 104 AA.
 AC Q9UL87;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Myosin-reactive immunoglobulin heavy chain variable region (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98277139; PubMed=9614934;
 RA Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,
 RA Young D.C.;
 RT "Myosin-reactive autoantibodies in rheumatic carditis and normal fetus."
 RL Clin. Immunol. Immunopathol. 87:184-192 (1998).
 DR EMBL; AF035027; AAD56263.1; --.
 DR HSSP; P01810; 2FBJ.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003596; Ig_v.
 DR Pfam; PF00047; Ig; 1.
 DR SMART; SM00406; IG; 1.
 DR PROSITE; PSS0835; IG_LIKE; 1.
 FT NON_TER 1
 FT NON_TER 104
 SQ SEQUENCE 104 AA; 11598 MW; 611DIA3F40B96E7B CRC64;

Query Match 58.8%; Score 50; DB 4; Length 104;
 Best Local Similarity 68.8%; Pred. No. 1.6;
 Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 2 ISENGRTINYADSVKG 17
 Db 30 ISWNGSGTYADSVKG 45

RESULT 4

Q8WU38 Q8WU38 PRELIMINARY; PRT; 573 AA.
 AC Q8WU38;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=tensil;
 RA Strausberg R.;
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC021276; AAH21276.1; --.
 DR PIR; S21205; S21205.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003596; Ig_v.
 DR Pfam; PF00047; Ig; 4.
 DR SMART; SM00406; IG; 1.
 DR PROSITE; PSS0835; IG_LIKE; 4.
 DR PROSITE; PSS0290; IG_MHC; 2.
 KW Hypothetical protein.

SQ SEQUENCE 573 AA; 62967 MW; FD072344033AC530 CRC64;

Query Match 58.8%; Score 50; DB 4; Length 573;
 Best Local Similarity 68.8%; Pred. No. 11;
 Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 2 ISENGRTINYADSVKG 17
 Db 70 ISWNGSGTYADSVKG 85

RESULT 5

Q9HCC1 Q9HCC1 PRELIMINARY; PRT; 112 AA.
 AC Q9HCC1;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Single chain Fv (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kikuchi M., Takeda C., Tsuchimoto Y., Asada S., Nagata K.;
 RT "An antibody fragment2A3 specific for native lysozyme : isolation from a human synthetic phage display library and characterization."
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB049915; BAB16829.1; --.
 DR HSSP; P01772; 2FB4.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003596; Ig_v.
 DR Pfam; PF00047; Ig; 1.
 DR SMART; SM00406; IG; 1.
 DR PROSITE; PSS0835; IG_LIKE; 1.
 FT NON_TER 1
 FT NON_TER 112
 SQ SEQUENCE 112 AA; 12243 MW; 24F1A45EC3B84788 CRC64;

Query Match 56.5%; Score 48; DB 4; Length 112;
 Best Local Similarity 62.5%; Pred. No. 3.7;
 Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 2 ISENGRTINYADSVKG 17
 Db 51 INWNGSTGYADSVKG 66

RESULT 6

Q9UL93 Q9UL93 PRELIMINARY; PRT; 116 AA.
 AC Q9UL93;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Myosin-reactive immunoglobulin heavy chain variable region (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98277139; PubMed=9614934;
 RA Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,
 RA Young D.C.;
 RT "Myosin-reactive autoantibodies in rheumatic carditis and normal fetus."
 RL Clin. Immunol. Immunopathol. 87:184-192 (1998).
 DR EMBL; AF035021; AAD56257.1; --.
 DR PIR; P0120; P0120.
 DR HSSP; P01772; 2FB4.

DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00835; IG_LIKE; 1.
FT NON_TER 116 116
SQ SEQUENCE 116 AA; 12434 MW; 0DA0348154DD6061 CRC64;

Query Match 55.3%; Score 47; DB 4; Length 116;
Best Local Similarity 64.7%; Pred. No. 5.7;
Matches 11; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 VISENGRTINYADSVKG 17
|||:|:|||||
Db 49 VISYDGSNKYYADSVKG 65

RESULT 7

Q8KW71 Q8KW71 PRELIMINARY; PRT; 341 AA.
AC Q8KW71; 2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE RC119.
OS Ruegeria sp. Prib.
OG Plasmid pSD25.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
OC Rhodobacteraceae; Ruegeria.
OX NCBI_TaxID=185588;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Prib;
RA Zhong Z.; Toudkarian A.; Helinski D.R.;
RT "Nucleotide sequence-based characterization of two cryptic plasmids
from the marine bacterium Ruegeria isolate Prib.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF416331; AAN05192.1; -.
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
KW Plasmid.
SQ SEQUENCE 341 AA; 37975 MW; 960D96A3D851D833 CRC64;

Query Match 55.3%; Score 47; DB 2; Length 341;
Best Local Similarity 66.7%; Pred. No. 19;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 6 GRTINYADSVKG 17
|:|||||:|
Db 277 GATINYADATSG 288

RESULT 8

Q8WUK1 Q8WUK1 PRELIMINARY; PRT; 613 AA.
AC Q8WUK1;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=tonsil;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC020240; AAN0240.1; -.
DR PIR; P0120; P0120.
DR PIR; S15590; S15590.
DR InterPro; IPR007110; Ig-like.

DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 5.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00835; IG_LIKE; 5.
DR PROSITE; PS00290; IG_MHC; 3.
KW Hypothetical protein.
SQ SEQUENCE 613 AA; 67296 MW; 60C7F5950671E315 CRC64;

Query Match 55.3%; Score 47; DB 4; Length 613;
Best Local Similarity 64.7%; Pred. No. 37;
Matches 11; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 VISENGRTINYADSVKG 17
|||:|:|||||
Db 69 VISYDGSNKYYADSVKG 85

RESULT 9

Q9UL71 Q9UL71 PRELIMINARY; PRT; 121 AA.
AC Q9UL71;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Myosin-reactive immunoglobulin heavy chain variable region
(Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RA Wu X.; Liu B.; Van der Merwe P.L.; Kallis N.N.; Berney S.M.;
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035043; AAD56279.1; -.
DR HSP; P01772; 2PB4.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00835; IG_LIKE; 1.
FT NON_TER 121 121
FT NON_TER 121 121
SQ SEQUENCE 121 AA; 13154 MW; 2F045CCFA5D50736 CRC64;

Query Match 54.1%; Score 46; DB 4; Length 121;
Best Local Similarity 58.8%; Pred. No. 8.8;
Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 VISENGRTINYADSVKG 17
|||:|:|||||
Db 50 LISDGGSTYYADSVKG 66

RESULT 10

Q9A8H8 Q9A8H8 PRELIMINARY; PRT; 466 AA.
AC Q9A8H8;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein CC1375.
GN CC1375.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;
OC Caulobacteraceae; Caulobacter.
OX NCBI_TaxID=155892;
RN [1]

```

DR Pfam; PF00034; cytochrome_c; 2.
DR PRINTS; PR00605; CYTCROME.C.
DR PROSITE; PS00190; CYTOCHROME_C; 3.
DR PROSITE; PS00030; RRM_RNP_1; 1.
KW Complete proteome.
SQ SEQUENCE 1187 AA; 127874 MW; A237F711DB6CE4B5 CRC64;

Query Match          54.1%; Score 46; DB 16; Length 1187;
Best Local Similarity 47.1%; Pred. No. 1.1e+02;
Matches      8; Conservative    6; Mismatches   3; Indels    0; Gaps    0;

Qy  1 VISENGRTINYADSVKG 17
Dd  152 VIASNGSTLSFAELVQG 168
|||: || |:::||: ||
RESULT 12
O26670 PRELIMINARY; PRT; 310 AA.
ID O26670 AC O26670;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein MTH570.
GN MTH570.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
OX NCBI_TaxID=187420;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Delta H;
RX MEDLINE=98037514; PubMed=9371463;
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lum W., Pothier B., Qiu D.,
RA Spadafora R., Vicare R., Wang Y., Wierzowski J., Gibson R.,
RA Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McDougall S., Shiner G., Goyal A., Pietrowski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT delah; functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7155(1997).
RW EMBL; AE000839; AAC85076.1; -.
DR PIR; E69175; E69175.
KW Hypochemical protein; Complete proteome.
SQ SEQUENCE 310 AA; 35223 MW; 70F30E12BA6110A6 CRC64;

Query Match          52.9%; Score 45; DB 17; Length 310;
Best Local Similarity 60.0%; Pred. No. 37;
Matches     9; Conservative    2; Mismatches   4; Indels    0; Gaps    0;

Qy  1 VISENGRTINYADSV 15
Dd  131 VISENLATINYEQI 145
||||| ||||| : :
RESULT 13
Q96BB9 PRELIMINARY; PRT; 597 AA.
ID Q96BB9 AC Q96BB9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypochemical protein.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=B-cell;
RA Strausberg R.;
```

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC015760; AALH5760.1; -
 DR InterPro; IPR007110; IG-like.
 DR InterPro; IPR003006; IG_MHC.
 DR InterPro; IPR003596; IG_v.
 DR Pfam; PF00047; IG; 5.
 DR SMART; SM00406; IGV; 1.
 DR PROSITE; PS00835; IG-LIKE; 5.
 DR PROSITE; PS00290; IG_MHC; 3.
 KW Hypothetical protein.
 SQ SEQUENCE 597 AA; 65039 MW; 4FCA3AD8BCE263D9 CRC64;

Query Match 52.9%; Score 45; DB 4; Length 597;
 Best Local Similarity 62.5%; Pred. No. 77;
 Matches 10; Conservative 2; Mismatches 4; Indels 0;

Qy 2 ISENGRTINVDVSKG 17

Db 70 ISGSGGSTRYADSVK 85

RESULT 14

Q9BLH9 PRELIMINARY; PRT; 640 AA.
 AC Q9BLH9
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE TRASSC4 protein (Fragment).
 GN TRASSC4
 OS Samia cynthia (Cynthia moth) (Ailanthus silkmoth).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Lepidoptera; Glossata; Bombycoidea;
 OC Saturniidae; Saturniinae; Attacini; Samia.
 OX NCBI_TaxID=7127;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TRANSPOSON=TRASSC4; TISSUE=silk gland;
 RX MEDLINE=21219196; PubMed=11319268;
 RA Kubo Y., Okazaki S., Anzai T., Fujiwara H.;
 RT "Structural and Phylogenetic Analysis of TRAS, Telomeric Repeat-
 Specific Non-LTR Retrotransposon Families in Lepidopteran Insects.";
 RL MOL. Biol. Evol. 18:848-857(2001).
 DR EMBL; AB046674; BAB21517.1; -
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006278; P:RNA dependent DNA replication; IEA.
 DR InterPro; IPR005135; Exo_endo_phos.
 DR InterPro; IPR000477; RVTse.
 DR Pfam; PF03372; Exo_endo_phos; 1.
 DR Pfam; PF00078; Ivt; 1.
 KW RNA-directed DNA polymerase; Transferase.
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 640 AA; 73783 MW; 3ADAB436AE58060C CRC64;

Query Match 52.9%; Score 45; DB 5; Length 640;
 Best Local Similarity 50.0%; Pred. No. 83;
 Matches 8; Conservative 5; Mismatches 3; Indels 0;

Qy 1 VISENGRTINVDVSK 16

Db 337 LLSQNGQLNPEDSVK 352

RESULT 15

Q88LF9 PRELIMINARY; PRT; 671 AA.
 AC Q88LF9
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Excinuclease ABC, B subunit.
 GN UVRB OR PP1974.
 OS Pseudomonas putida (strain KT2440).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=160488;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22423080; PubMed=12534463;
 RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
 RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
 RA Brinkac L., Beanan M., DeBoy R.T., Daugherty S., Kolonay J.,
 RA Madupu R., Nelson W., White O., Peterson J., Khouri H., Hance I.,
 RA Chris Lee P., Holtzapple E., Scanlan D., Tran K., Moazzaz A.,
 RA Uterback T., Rizzo M., Lee K., Kosack D., Moestl D., Wedler H.,
 RA Lauber J., Stjepandic D., Hoheisel J., Straetz M., Helm S.,
 RA Kiewitz C., Eisen J., Timmis K.N., Duesterhoft A., Tuemmler B.,
 RA Fraser C.M.;
 RT "Complete genome sequence and comparative analysis of the
 RT metabolically versatile Pseudomonas putida KT2440.";
 RL Environ. Microbiol. 4:799-808(2002).
 DR EMBL; AE016781; AAN67589.1; -
 DR TIGR; PP1974; -
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0008026; F:ATP dependent helicase activity; IEA.
 DR GO; GO:0003877; F:DNA binding; IEA.
 DR GO; GO:0004518; F:nuclease activity; IEA.
 DR GO; GO:0006289; P:nucleotide-excision repair; IEA.
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR001650; Helicase_C.
 DR InterPro; IPR001943; UvrB/C_1.
 DR Pfam; PF00271; helicase_C; 1.
 DR Pfam; PF02151; Uvr; 1.
 DR PROSITE; PS0151; UVR; 1.
 KW Complete proteome.
 SQ SEQUENCE 671 AA; 76034 MW; 5175F5EBDBF41123 CRC64;

Query Match 52.9%; Score 45; DB 16; Length 671;
 Best Local Similarity 69.2%; Pred. No. 88;
 Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 5 NGRTINVDVSKG 17

Db 545 NGRALLYADNVG 557

RESULT 16

Q8YC77 PRELIMINARY; PRT; 729 AA.
 AC Q8YC77
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Alkaline phosphatase [EC 3.1.3.1].
 GN BMEI0655
 OS Brucella melitensis.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Brucellaceae; Brucella.
 OX NCBI_TaxID=29459;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=16M / ATCC 23456 / Biotype 1;
 RX MEDLINE=20020109; PubMed=11756688;
 RA DelVecchio V.G., Kapral V., Redkar R.J., Patra G., Mujer C., Los T.,
 RA Ivanova N., Anderson I., Bhattacharya A., Lykidis A., Reznik G.,
 RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goldsman E.,
 RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
 RA Haselkorn R., Kyrpides N., Overbeek R.;
 RT "The genome sequence of the facultative intracellular pathogen
 RT Brucella melitensis.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
 DR EMBL; AE009700; AAL53897.1; -
 DR PIR; AF3591; AF3591.

DR GO; GO:0004035; F:alkaline phosphatase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
KW Hydrolase; Complete proteome.
SQ SEQUENCE 729 AA; 78103 MW; 21F7888632E30CDE CRC64;

Query Match 52.9%; Score 45; DB 16; Length 729;
Best Local Similarity 50.0%; Pred. No. 96;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 2 ISENGRTINYADSVKG 17

Db 57 VSEDKTLIYSDSPYG 72

RESULT 17

Q8FW49 PRELIMINARY; PRT; 729 AA.
AC Q8FW49;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Conserved hypothetical protein.
GN BRA0616.
OS Brucella suis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1330 / Biovar 1;
RX MEDLINE=22247741; PubMed=12271122;
RA Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,
Read T.D., Dodson R.J., Unayam L., Brinkac L.M., Beanan M.J.,
Daugherty S.C., Deboy R.T., Durkin A.S., Kolonay J.F., Madupu R.,
Nelson W.C., Ayodeji B., Kraul M., Shetty J., Malek J., Van Aken S.E.,
Riedmuller S., Tetrelin H., Gill S.R., White O., Salzberg S.L.,
Hoover D.L., Lindler L.E., Halling S.M., Boyle S.M., Fraser C.M.;
RT "The Brucella suis genome reveals fundamental similarities between
animal and plant pathogens and symbionts."
RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153 (2002).
DR EMBL; AE014558; AAN33805.1; -.
DR TIGR; BRA0616; -.
KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 729 AA; 78147 MW; 964E74BAP188B292 CRC64;
Query Match 52.9%; Score 45; DB 16; Length 729;
Best Local Similarity 50.0%; Pred. No. 96;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 2 ISENGRTINYADSVKG 17

Db 57 VSEDKTLIYSDSPYG 72

RESULT 18

Q9IV97 PRELIMINARY; PRT; 853 AA.
AC Q9IV97;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Env polyprotein.
GN Env.
OS Human immunodeficiency virus 1.
OC Viruses; Retroviridae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VII1035;
RX MEDLINE=20418941; PubMed=10954895;
RA Janssens W., Salminen M.O., Laukkanen T., Heyndrickx L.,
Van der Auwera G., Colebunders R., McCutchan F.E., Van der Groen G.;
RT "Near full-length genome analysis of HIV type 1 CRF02_AG subtype C and

RT CRF02_AG subtype G recombinants."
RL AIDS Res. Hum. Retroviruses 16:1183-1189 (2000).
DR EMBL; AJ276595; CAB82224.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 853 AA; 95995 MW; 85E4A320F516B1B2 CRC64;

Query Match 52.9%; Score 45; DB 15; Length 853;
Best Local Similarity 50.0%; Pred. No. 1.1e+02;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 2 ISENGRTINYADSVKG 17

Db 136 INSGTAINITESIKG 151

RESULT 19

Q9A430 PRELIMINARY; PRT; 1105 AA.
AC Q9A430;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE TonB-dependent receptor.
GN CC3013.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacteriales;
OC Caulobacteraceae; Caulobacter.
OX NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647;
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
Radocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
Uterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus."
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141 (2001).
DR EMBL; AE005964; AAK24975.1; -.
DR PIR; C87622; C87622.
DR TIGR; CC3013; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000531; TonB_boxC.
DR Pfam; PF00593; TonB_dep_Rec; 1.
KW Receptor; Complete proteome.

SQ SEQUENCE 1105 AA; 121226 MW; C921159A608AEB99 CRC64;
Query Match 52.9%; Score 45; DB 16; Length 1105;
Best Local Similarity 57.1%; Pred. No. 1.5e+02;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 3 SENGRTINYADSVK 16

Db 590 SRDGRVTNVAKGTK 603

RESULT 20

Q9ZQX5 PRELIMINARY; PRT; 3449 AA.
ID Q9ZQX5

```
AC Q0ZQX5;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative transport protein.
GN T10P11.5 OR AT4G02650.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Kaplan N., Johnson D., Schutz K., Gnoj L., Hoffman J., Till S.,
RA de la Bastide M., Granat S., Hameed A., Gottesman T., Hasegawa A.,
RA Shohdy N., Parnell L., Dedhia N., Johnson A.F., Lodhi M.,
RA Martensen R., Chen E.Y., Wilson R., McCombie W.R.;
RT "Sequence of A. thaliana BAC T10P11 from chromosome IV.";
RN Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RA Spiegel L.A., Huang E.N., Nascimento L.U., de la Bastide M., Vil D.M.,
RA Preston R.R., Matero A., Shah R., O'Shaughnessy A., Rodriguez M.,
RA Shekher M., Schutz K., See L.H., Swaby I., Habermann K., Dedhia N.N.,
RA Mewes H.W., Lemcke K., Mayer K.F.X.;
RN Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[3]
RN SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC002330; AAC78268.1; -.
DR EMBL; AL161495; CAB77751.1; -.
DR PIR; T01083; T01083.
DR InterPro; IPR008938; ARM.
DR InterPro; IPR000409; Beige_BEACH.
DR InterPro; IPR001680; WD40.
DR Pfam; PF02138; Beach; 1.
DR Pfam; PF00400; WD40; 2.
DR Pfam; PF007848; Beige_BEACH; 1.
DR ProDom; PD007848; Beige_BEACH; 1.
DR SMART; SM00320; WD40; 2.
DR PROSITE; PS0197; BEACH; 1.
DR PROSITE; PS00678; WD_REPEATS_1; 1.
DR PROSITE; PS0082; WD_REPEATS_2; 1.
DR PROSITE; PS0294; WD_REPEATS_REGION; 1.
KW Hypothetical protein; Repeat; WD repeat.
SQ SEQUENCE 3449 AA; 385006 MW; 6BB0DBA778690184 CRC64;

Query Match 52.9%; Score 45; DB 10; Length 3449;
Best Local Similarity 47.1%; Pred. No. 5.5e+02;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 VISENGRTINYADSVKG 17
Db 185 VASEGKTVYEEKVEG 201

RESULT 21
Q9UL91 PRELIMINARY; PRT; 118 AA.
ID Q9UL91
AC Q9UL91
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Myosin-reactive immunoglobulin heavy chain variable region
DE (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RL MEDLINE=98277139; PubMed=9614934;
RX VISENGRTINYADSVKG 17
Q9UL91 PRELIMINARY; PRT; 118 AA.
ID Q9UL91
AC Q9UL91
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Myosin-reactive immunoglobulin heavy chain variable region
DE (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RL MEDLINE=98277139; PubMed=9614934;
RX VASEGKTVYEEKVEG 201
Q9UL91 PRELIMINARY; PRT; 118 AA.
ID Q9UL91
AC Q9UL91
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Myosin-reactive immunoglobulin heavy chain variable region
DE (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RL MEDLINE=98277139; PubMed=9614934;
RX VASEGKTVYEEKVEG 201
```

```
RA Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035023; AAD56259.1; -.
DR PIR; S21205; S21205.
DR HSP; P01772; 2FB4.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
FT NON_TER 118
FT NON_TER 118
SQ SEQUENCE 118 AA; 12843 MW; D0633949F2AC149D CRC64;

Query Match 51.8%; Score 44; DB 4; Length 118;
Best Local Similarity 68.8%; Pred. No. 18;
Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 ISENGRTINYADSVKG 17
Db 51 ISSTIITIVYADSVKG 66

RESULT 22
Q9UL84 PRELIMINARY; PRT; 122 AA.
ID Q9UL84
AC Q9UL84
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Myosin-reactive immunoglobulin heavy chain variable region
DE (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RL MEDLINE=98277139; PubMed=9614934;
RX Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035030; AAD56266.1; -.
DR HSP; P01772; 2FB4.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
FT NON_TER 122
FT NON_TER 122
SQ SEQUENCE 122 AA; 13579 MW; 36054D41366545B8 CRC64;

Query Match 51.8%; Score 44; DB 4; Length 122;
Best Local Similarity 62.5%; Pred. No. 19;
Matches 10; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 2 ISENGRTINYADSVKG 17
Db 51 ISNDGSKFYADSVKG 66

RESULT 23
Q9JH96 PRELIMINARY; PRT; 220 AA.
ID Q9JH96
AC Q9JH96
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
```

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Family 45 cellulase homologue.
OS unclassified eukaryotes.

OC Eukaryota.
OX NCBI TaxID=42452;

RN	[1]	SEQUENCE FROM N.A.
RP		

RA Ohtoko K., Ohkuma
RT "Diverse genes of

RT protists in the hindgut of termite *Reticulitermes speratus*.¹; RL Submitted (JUN-2000) to the ENBL/GenBank/DBJ databases.

DR EMBL; AB045165; BAA98035.1; -
DR HSSP; P43316; 2FNG.

DR GO:0005975: P:carbohydrate metabolism

DR GO;0003973; F:Carbonylurate Metab
DR InterPro; IPR009009; Barwin like.
DR InterPro; IPR0000334; Glyco Hydrol 45

DR InterPro; IPR000334; Glyco_hydro_45.
DR Pfam; PF02015; Glyco_hydro_45; 1.
DR PROSITE; PS01140; GLYCOSYL_TRANSFERASE; 1.

DR PROSITE; PS011140; GLYCOSYL HYDROL_F45; 1.
SQ SEQUENCE 220 AA; 23148 MW; 4F0652F8E8D

Query Match	Score 44;	DB 14;	Length
51.8%			

Best Local Similarity 52.9%; Pred. No. 37;
Matches 9; Conservative 1; Mismatches 7; Indels

QY	
1 VISENGRTINYADSVKG	17

13 VFGEGRTRYWDCKG 29

RESULT 24
Q91Z05

Q91205	PRELIMINARY;	PRT;	473 AA.
ID Q91205			
AC Q91205:			

DT	Q1203,	01-DEC-2001 (TREMBlrel. 19, Created)	01-DEC-2001 (TREMBlrel. 19, Last sec
DT			
DT			

DI 01-DEC-2001 (EMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein

DE	Hypothetical protein.
GN	AU044919.
OS	<i>Mus musculus</i> (Mouse)

US *Mus musculus* (Mouse).
OC Eukaryota; Metazoa; Chordata;
OC Mammalia; Eutheria; Rodentia;

OC Mammalia; Eutheria; Rodentia;
OX NCBI_TaxID=10090;

RP SEQUENCE FROM N.A.

RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC010327; AAH10327.1; -.
DR MGD; MGI:2144967; AU044919.

DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.

DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR007110; Ig-like.

```

DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_v.

```

DR PFam: PF00047; ig: 3.
DR SMART: SM00406; Igv: 1.

DR PROSITE; PS00190; CYTOCHROME_C; 1.
DR PROSITE; PS50A35; IG LIKE; 4.

DR PROSITE; PS00290; IG_MHC; 1.
KW Homothetical protein

NW hypochetrical protein.
SQ SEQUENCE 473 AA; 51946 MW; CF625F008932AF12 CRC64;

Query Match 51.8%; Score 44; DB 11; Length 473;

Best Local Similarity 56.2%; Pred. No. 87;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVKG 17

Db 70 INSGSTIIYYADTVKG 85

RESULT 25

Q7VJ23	PRELIMINARY;	PRT;	504 AA.
ID	Q7VJ23		
AC	Q7VJ23;		
DT	01-OCT-2003 (TrEMBLrel. 25, Created)		
DT	01-OCT-2003 (TrEMBLrel. 25, Last sequence update)		
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)		
DE	FOF1-type ATP synthase (EC 3.6.3.14).		
GN	ATPA OR HH0427.		
OS	Helicobacter hepaticus.		
OC	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;		
OC	Helicobacteraceae; Helicobacter.		
ON	NCBI_TaxID=32025;		
OX	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=ATCC 51449 / 3B1;		
RX	MEDLINE=22709201; PubMed=12810954;		
RA	Suierbaum S., Josephans C., Sterzenbach T., Drescher B., Brandt P.,		
RA	Ball M., Droege M., Fartmann B., Fischer H.-P., Ge Z., Hoerster A.,		
RA	Holland R., Klein K., Koenig J., Macko L., Mendz G.L., Nyakatura G.		
RA	Schauer D.B., Shen Z., Weber J., Froesch M., Fox J.G.;		
RT	"The complete genome sequence of the carcinogenic bacterium		
RT	Helicobacter hepaticus."		
RL	Proc. Natl. Acad. Sci. U.S.A. 100:7901-7906 (2003).		
DR	ENBL; AE017145; AAP77024.1; -.		
SK	Hydrolase; Complete proteome.		
QW	SEQUENCE 504 AA; 55287 MW; 88B831A9C5FF2469 CRC64;		

Query Match 51.8%; Score 44; DB 16; Length 504;

Best Local Similarity	64.3%	Pred. No. 94	
Matches	9	Conservative	1: Mismatches 4: Indels 0: Gaps 0:

OV 2 ISENGRTINYADSV 15

27 ISETGKVIAYADGV 40

Search completed: September 24, 2004, 02:18:47
Job time : 88.4717 secs

000000 / T / 1 : 0000

This page Blank (uspto)

RESULT 2	
Q8NSK4	
ID	Q8NSK4
AC	PRELIMINARY; PRT; 499 AA.
DT	Q8NSK4;
DT	01-OCT-2002 (TrEMBLrel. 22, Created)
DT	01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE	01-WAR-2004 (TrEMBLrel. 26, Last annotation update)
DE	MGC27165 protein.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

RT "Generation and initial analysis of more than 15,000 full-length human
 RL and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=Spleen;
 RA Klein S., Strausberg R.;
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC077477; AAH77477.1; --

DR InterPro; IPR003599; IG.

DR InterPro; IPR007110; IG-like.

DR InterPro; IPR003597; IG.C1.

DR InterPro; IPR003006; IG.MHC.

DR InterPro; IPR003596; IG.V.

DR InterPro; IPR001680; WD40.

DR Pfam; PF07654; Cl-set; 4.

DR Pfam; PF00047; ig; 3.

DR SMART; SM00409; IG; 4.

DR SMART; SM00407; IGc1; 4.

DR SMART; SM00406; IG; 1.

DR PROSITE; PS00835; IG LIKE; 5.

DR PROSITE; PS00290; IG.MHC; UNKNOWN 4.

DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN 1.

SQ SEQUENCE 614 AA; 68253 MW; 2631B7CF955270C0 CRC64;

Query Match 60.0%; Score 51; DB 2; Length 614;

Best Local Similarity 64.7%; Pred. No. 12;

Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 VISENGRTINYADSVKG 17

||||:|||||

Db 67 VISSDGGTYADSVKG 83

RESULT 5

Q9UL87 PRELIMINARY; PRT; 104 AA.

AC Q9UL87;

DT 01-MAY-2000 (TREMELrel. 13, Created)

DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)

DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)

DE Myosin-reactive immunoglobulin heavy chain variable region

DE (Fragment).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=98277139; PubMed=9614934; DOI=10.1006/clin.1998.4531;

RA Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,

RA Young D.C.;

RT "Myosin-reactive autoantibodies in rheumatic carditis and normal

RT fetus";

RL Clin. Immunol. Immunopathol. 87:184-192(1998).

DR EMBL; AF035027; AAD56263.1; --

DR HSP; P01810; 2FBJ.

DR InterPro; IPR007110; IG-like.

DR InterPro; IPR003596; IG.V.

DR SMART; SM00406; IG; 1.

DR PROSITE; PS00835; IG LIKE; 1.

FT NON_TER 1 104

SQ SEQUENCE 104 AA; 11598 MW; 611D1A3F40E96E7B CRC64;

Query Match 58.8%; Score 50; DB 2; Length 104;

Best Local Similarity 68.8%; Pred. No. 2.6;

Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVKG 17

||||:|||||

Db 30 ISWNSGSGIYADSVKG 45

RESULT 6

Q8WU38

ID Q8WU38 PRELIMINARY; PRT; 573 AA.

AC Q8WU38;

DT 01-MAR-2002 (TREMELrel. 20, Created)

DT 01-MAR-2002 (TREMELrel. 20, Last sequence update)

DT 01-MAR-2004 (TREMELrel. 26, Last annotation update)

DE Hypothetical protein.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Primary B-Cells;

RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,

RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,

RA Jones S.J., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Primary B-Cells;

RA Strausberg R.;

RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC021276; AAH21276.1; --

DR PIR; S21205; S21205.

DR PIR; S30532; S30532.

DR HSP; P18529; 118K.

DR Pfam; PF07654; Cl-set; 2.

DR SMART; SM00407; IG; 1.

DR SMART; SM00406; IG; 1.

DR PROSITE; PS50835; IG LIKE; 4.

DR PROSITE; PS00290; IG.MHC; UNKNOWN_2.

KW Hypothetical protein.

SQ SEQUENCE 573 AA; 62967 MW; FD072344033AC530 CRC64;

Query Match 58.8%; Score 50; DB 2; Length 573;

Best Local Similarity 68.8%; Pred. No. 16;

Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

RESULT 7

HV3U HUMAN

ID HV3U HUMAN STANDARD; PRT; 120 AA.

AC P01782;

DT 21-JUL-1986 (Rel. 01, Created)

DT 21-JUL-1986 (Rel. 01, Last sequence update)

DT 05-JUL-2004 (Rel. 44, Last annotation update)

DE Ig heavy chain V-III region DOB.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=80020921; PubMed=114209;
RA Steiner L.A., Garcia Pardo A., Margolies M.N.;
RA "Amino acid sequence of the heavy-chain variable region of the
RT crystallizable human myeloma protein Dob.";
RL Biochemistry 18:4068-4080(1979).
RN [2]
RP CRYSTALLIZATION.
RX MEDLINE=80020920; PubMed=114208;
RA Steiner L.A., Lopes A.D.;
RA "The crystallizable human myeloma protein Dob has a hinge-region
RT deletion.";
RL Biochemistry 18:4054-4067(1979).
RN [3]
RP MISCELLANEOUS: This gamma-1 myeloma protein has a deletion in the
CC hinge region. There are no light-heavy or inter-heavy chain
CC disulfide bonds.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
DR PIR; A90431; GIHUBD.
DR HSSP; P01772; 2FB4.
DR GO; GO:0005576; C:extracellular; NAS.
DR GO; GO:0003823; F:antigen binding; NAS.
DR GO; GO:0006955; P:immune response; NAS.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG LIKE; 1.
KW Direct protein sequencing; Immunoglobulin V region.
FT DOMAIN 1 112 Ig-like.
FT NON_TER 120 120
SQ SEQUENCE 120 AA; 13440 MW; 880DDE307C4B2627 CRC64;

Query Match 57.6%; Score 49; DB 1; Length 120;
Best Local Similarity 62.5%; Pred. No. 4.4;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVKG 17
| : : : : |||||
Db 51 ITWNGGSLYADSVKG 66

RESULT 8
HV3A_HUMAN STANDARD; PRT; 122 AA.
AC P01762;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Ig heavy chain V-III region TRO.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE (MYELOMA PROTEIN TRO).
RX MEDLINE=76023781; PubMed=809331;
RA Kratzin H., Altevogt P., Ruban E., Kortt A., Starosck K.,
RA Hilschmann N.;
RT "The primary structure of a monoclonal IgA-immunoglobulin (IgA Tro.),
RT II. The amino acid sequence of the H-chain, alpha-type, subgroup III;
RT structure of the complete IgA-molecule.";
RL Hoppe-Seyler's Z. Physiol. Chem. 356:1337-1342(1975).
CC -1- MISCELLANEOUS: The sequence of the C region is also given.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
DR PIR; A02045; ALHUTR.
DR HSSP; P01783; 1IGC.
DR GO; GO:0005576; C:extracellular; NAS.
DR GO; GO:0003823; F:antigen binding; NAS.
DR GO; GO:0006955; P:immune response; NAS.
DR InterPro; IPR007110; Ig-like.

DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; IGV; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG LIKE; 1.
KW Direct protein sequencing; Immunoglobulin V region;
KW Pyrolydione carboxylic acid.
FT DOMAIN 1 108 Ig-like.
FT MOD_RES 1 122 Pyrrolidone carboxylic acid.
FT NON_TER 122 122
SQ SEQUENCE 122 AA; 13472 MW; 2E21A11DA04D80F9 CRC64;

Query Match 57.6%; Score 49; DB 1; Length 122;
Best Local Similarity 62.5%; Pred. No. 4.4;
Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVKG 17
| : : : : |||||
Db 51 IGGSGSTLYADSVKG 66

RESULT 9
Q9HCC1 PRELIMINARY; PRT; 112 AA.
AC Q9HCC1;
DT 01-MAR-2001 (TREMELrel. 16, Created)
DT 01-MAR-2001 (TREMELrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Single chain Fv (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kikuchi M., Takeda C., Tsujimoto Y., Asada S., Nagata K.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB049915; BAB16829.1; --
DR HSSP; P01783; 1IGC.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003596; Ig_v.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG LIKE; 1.
FT NON_TER 1 112
FT NON_TER 112 112
SQ SEQUENCE 112 AA; 12243 MW; 24F1A45EC3B84788 CRC64;

Query Match 56.5%; Score 48; DB 2; Length 112;
Best Local Similarity 62.5%; Pred. No. 5.9;
Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 ISENGRTINYADSVKG 17
| : : : : |||||
Db 51 INWNGSGTGYADSVKG 66

RESULT 10
HV16_MOUSE STANDARD; PRT; 136 AA.
AC P01783;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Ig heavy chain V region MOPC 21 precursor (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81234548; PubMed=6788376; DOI=10.1016/0092-8674(81)90089-1;
RA Bothwell A.L.M., Paskind M., Reth M., Imanishi-Kari T., Rajewsky K.,
RA Baltimore D.;
RT "Heavy chain variable region contribution to the Npb family of

antibodies: somatic mutation evident in a gamma 2a variable region.";
Cell 24:625-637(1981).
(2)
SEQUENCE OF 17-136.
MEDLINE=77100368; PubMed=401950;
Adetubo K., Milstein C., Secher D.S.;
"Molecular analysis of spontaneous somatic mutants.";
Nature 265:299-304(1977).

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; J00522; AAD15290.1; -;
PIR; E90809; GIM521.
PDB; 1IGC; X-ray; H=-;
InterPro; IPR007110; Ig-like.
InterPro; IPR003596; Ig_v.
Pfam; PF00047; Ig; 1.
SMART; SM00406; IGV; 1.
PROSITE; PS00835; IG LIKE; 1.
3D-structure; Direct protein sequencing; Immunoglobulin V region;
Signal.
NON TER 1 1
SIGNAL <1 16
CHAIN 17 136
DOMAIN 115 119
DOMAIN 120 136
DISULFID 38 112
CONFLICT 75 78
CONFLICT 89 90
CONFLICT 115 115
CONFLICT 120 120
STRAND 19 23
STRAND 26 28
TURN 30 31
TURN 34 41
HELEX 45 47
STRAND 50 55
TURN 61 67
STRAND 69 70
STRAND 74 76
HELEX 78 80
STRAND 81 81
TURN 82 83
STRAND 84 89
TURN 90 93
STRAND 94 99
HELEX 104 106
TURN 108 114
STRAND 125 126
STRAND 130 134
NON TER 136 136
SEQUENCE 136 AA; 15071 MW; 2276A98DBDBF7016 CRC64;

Query Match 56.5%; Score 48; DB 1; Length 136;
Best Local Similarity 56.2%; Pred. No. 7.3;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 ISENGRTINADVSKG 17
DB 67 ISSGSSTLRVADTVKG 82

RESULT 11
Q6N089
ID Q6N089 PRELIMINARY; PRT; 472 AA.
AC Q6N089;

05-JUL-2004 (TrEMBLrel. 27, Created)
05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
Hypothetical protein DXFZP686P15220.
GN Name=DXFZP686P15220;
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human rectum tumor;
RG The German Human cDNA Consortium;
RA Wambutt R., Heubner D., Mewes H.W., Weil B., Amid C., Oeanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX640627; CAB45781.1; -;
DR HSSP; P01861; IADQ.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig ci.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003596; IG_v.
DR Pfam; PF07654; Cl-set; 3.
DR SMART; SM00409; IG; 2.
DR SMART; SM00407; IGV; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00835; IG LIKE; 4.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_2.
KW Hypothetical protein.
SQ SEQUENCE 472 AA; 51724 MW; 26CB340D0046D279 CRC64;

Query Match 56.5%; Score 48; DB 2; Length 472;
Best Local Similarity 68.8%; Pred. No. 28;
Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 ISENGRTINADVSKG 17
DB 70 ISWNGSIAVADSVKG 85

RESULT 12
Q6GMX2
ID Q6GMX2 PRELIMINARY; PRT; 493 AA.
AC Q6GMX2
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Splice;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Basak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.E., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,

RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.

RC TISSUE=Spleen;
RA Strausberg R.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR ENBL; BC073771; AH73771.1; -
DR InterPro; IPR003599; Ig-like.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig-G1.
DR InterPro; IPR0031006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF07654; C1-set; 2.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00409; IG; 4.
DR SMART; SM00407; IG1; 2.
DR SMART; SM00406; IG1; 1.
DR PROSITE; PS0835; IG LIKE; 4.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 493 AA; 52865 MW; 55B999305B286203 CRC64;

Query Match 56.5%; Score 48; DB 2; Length 493;
Best Local Similarity 56.2%; Pred. No. 30;
Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 2 ISENGRTINYADSVKG 17
Db 70 INSDGSTSYADSVKG 85

RESULT 13
HV3T HUMAN STANDARD; PRT; 116 AA.
ID HV3T HUMAN STANDARD; PRT; 116 AA.
AC P01781;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Ig heavy chain V-III region GAL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=75059123; PubMed=4803843;
RA Watanabe S., Barnikol H.U., Horn J., Bertram J., Hilschmann N.;
RT "The primary structure of a monoclonal IgM-immunoglobulin
RT (macroglobulin Gal.), II: the amino acid sequence of the H-chain (mu-
RT type), subgroup H III. Architecture of the complete IgM-molecule.";
RL Hoppe-Seyler's Z. Physiol. Chem. 354:1505-1509(1973).
RN [2]

REVISION TO 28-33.
RA Hilschmann N.;
RL Submitted (JUN-1975) to the PIR data bank.
CC -1- MISCELLANEOUS: This mu chain was isolated from a Waldenstrom's
CC macroglobulin.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
DR PIR; A02064; M3HUGL.
DR HSP; P01772; 2F84.
DR GO; GO:0005576; C:extracellular; NAS.
DR GO; GO:0003823; F:antigen binding; NAS.
DR GO; GO:0006955; P:immune response; NAS.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IG1; 1.
DR PROSITE; PS0835; IG LIKE; 1.
KW Direct protein sequencing; Immunoglobulin V region.
DOMAIN 1 112 Ig-like.

FT NON_TER 116 116
SQ SEQUENCE 116 AA; 12730 MW; 2C67CA9A9AAA1282 CRC64;
Query Match 55.3%; Score 47; DB 1; Length 116;
Best Local Similarity 50.0%; Pred. No. 9;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 2 ISENGRTINYADSVKG 17
Db 51 IKZBGSGZBYVDVSKG 66

RESULT 14
Q9UL93 PRELIMINARY; PRT; 116 AA.
ID Q9UL93
AC Q9UL93;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Myosin-reactive immunoglobulin heavy chain variable region
DE (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934; DOI=10.1006/clin.1998.4531;
RA Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.;
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035021; AAD56257.1; -
DR PIR; PH1644; PH1644.
DR PIR; P01720; FL0120.
DR HSP; P01772; 2F84.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003596; Ig_v.
DR SMART; SM00406; IG1; 1.
DR PROSITE; PS0835; IG LIKE; 1.
FT NON_TER 116 116
FT NON_TER 116 116
SQ SEQUENCE 116 AA; 12434 MW; 0DA0348154DD6061 CRC64;

Query Match 55.3%; Score 47; DB 2; Length 116;
Best Local Similarity 64.7%; Pred. No. 9;
Matches 11; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 VISENGRTINYADSVKG 17
Db 49 VISYDGSNKYADSVKG 65

RESULT 15
Q65ZC9 PRELIMINARY; PRT; 240 AA.
ID Q65ZC9
AC Q65ZC9;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Single-chain Fv (Fragment).
GN Name=scfv;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C1g/7;
RX MEDLINE=97362799; PubMed=9219263;
RA Kontermann R.E., Wing M.G., Winter G.;
RT "Complement recruitment using bispecific diabodies.";

Query Match 55.3%; Score 47; DB 2; Length 116;
Best Local Similarity 64.7%; Pred. No. 9;
Matches 11; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 VISENGRTINYADSVKG 17
Db 49 VISYDGSNKYADSVKG 65

RESULT 15
Q65ZC9 PRELIMINARY; PRT; 240 AA.
ID Q65ZC9
AC Q65ZC9;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Single-chain Fv (Fragment).
GN Name=scfv;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C1g/7;
RX MEDLINE=97362799; PubMed=9219263;
RA Kontermann R.E., Wing M.G., Winter G.;
RT "Complement recruitment using bispecific diabodies.";

```

RL Nat. Biotechnol. 15:629-631(1997).
DR EMEL; Y13056; CAA73499.1; -.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003596; IG_v.
DR Pfam; PF00047; ig; 2.
DR SMART; SM00409; IG; 2.
DR SMART; SM00406; IGv; 2.
DR PROSITE; PS00835; IG LIKE; 2.
FT NON TER 1
FT NON TER 240
SQ SEQUENCE 240 AA; 25569 MW; FDCFD3645F64B373 CRC64;

Query Match 55.3%; Score 47; DB 2; Length 240;
Best Local Similarity 64.7%; Pred.No. 20;
Matches 11; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 VISENGRTINYADSVKG 17
   |||:| |||||
Db 50 VISYDGSNKYADSVKG 66

```

Search completed: April 28, 2005, 18:24:27
Job time : 113.969 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 01:32:26 ; Search time 83.6415 Seconds
(without alignments)
43.915 Million cell updates/sec

Title: US-10-088-639A-2_COPY_226_238

Perfect score: 69

Sequence: 1 EGGPGTTSNRLDA 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database :

A_Geneseq_29Jan04:*

1: Geneseq1980s:*

2: Geneseq1990s:*

3: Geneseq2000s:*

4: Geneseq2001s:*

5: Geneseq2002s:*

6: Geneseq2003as:*

7: Geneseq2003bs:*

8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	DB ID	Description
1	69	100.0	249	4	AAB68087
2	44	63.8	527	4	AAU64492
3	44	63.8	527	6	ABM64701
4	44	63.8	527	6	ABM61011
5	43	62.3	749	4	ABM60469
6	41	59.4	19938	6	ABP76682
7	40	58.0	224	2	AAU74086
8	39	56.5	77	4	AAU85724
9	39	56.5	130	6	ABU41557
10	39	56.5	219	5	AAU83139
11	39	56.5	296	5	ABP43778
12	39	56.5	302	4	ABM68531
13	39	56.5	396	4	AAU94593
14	39	56.5	396	4	AAU67257
15	39	56.5	454	4	AAU35107
16	39	56.5	454	6	ABU29209
17	39	56.5	600	4	AAU85503
18	39	56.5	600	5	AAE24147
19	39	56.5	609	3	AAU85459
20	39	56.5	1157	4	ABM63351
21	39	56.5	1253	4	ABM60494
22	38	55.1	102	3	AAU59078
23	38	55.1	134	5	ABG76902
24	38	55.1	174	3	AAG24538
25	38	55.1	184	3	AAG24537

26	38	55.1	202	3	AAG24536	Aag24536 Arabidops
27	38	55.1	233	3	AAG47562	Aag47562 Arabidops
28	38	55.1	236	3	AAG47561	Aag47561 Arabidops
29	38	55.1	240	3	AAG47585	Aag47585 Arabidops
30	38	55.1	246	3	AAG47560	Aag47560 Arabidops
31	38	55.1	250	3	AAG47584	Aag47584 Arabidops
32	38	55.1	268	3	AAG47583	Aag47583 Arabidops
33	38	55.1	508	4	AAU44929	Aau44929 Propionib
34	38	55.1	508	4	ABM41448	Abm41448 Propionib
35	38	55.1	803	7	ADB70254	Adb70254 C. neofo
36	37	53.6	36	6	ABP78016	Abp78016 N. gonorr
37	37	53.6	56	4	ABM40703	Abm40703 Peptide #
38	37	53.6	56	4	AAU34466	Aau34466 Peptide #
39	37	53.6	56	4	ABM74353	Abm74353 Human bon
40	37	53.6	56	4	ABG56151	Abg56151 Human liv
41	37	53.6	58	4	AAU58303	Aau58303 Human bra
42	37	53.6	58	4	ABG52488	Abg52488 Human liv
43	37	53.6	60	4	ABM15753	Abm15753 Human ner
44	37	53.6	100	4	AAU64316	Aau64316 Propionib
45	37	53.6	100	6	ABM60835	Abm60835 Propionib
46	37	53.6	146	4	ABM12241	Abm12241 Human Zn
47	37	53.6	192	6	ABU39922	Abu39922 Protein e
48	37	53.6	198	3	AAU75960	Aau75960 Rat skin
49	37	53.6	198	4	AAU55899	Aau55899 Skin cell
50	37	53.6	198	5	ABM72099	Abm72099 Rat prote
51	37	53.6	285	3	AAU76103	Aau76103 Rat skin
52	37	53.6	285	3	AAU54321	Aau54321 A polypep
53	37	53.6	285	3	AAU37344	Aau37344 Murine AC
54	37	53.6	285	3	AAU30232	Aau30232 Human adi
55	37	53.6	285	4	AAU56042	Aau56042 Skin cell
56	37	53.6	285	4	AAU50374	Aau50374 Human adi
57	37	53.6	285	5	ABM72242	Abm72242 Rat prote
58	37	53.6	285	6	ABP70917	Abp70917 ZADJ-2 pr
59	37	53.6	285	6	ABM42037	Abm42037 Human ZAD
60	37	53.6	285	6	ABP70907	Abp70907 ZADJ-2 pr
61	37	53.6	285	6	AAU16572	Aau16572 ZADJ-2 se
62	37	53.6	285	6	ABR39802	AbR39802 ZADJ-2 po
63	37	53.6	285	6	ABR40090	AbR40090 ZADJ-2 pr
64	37	53.6	285	6	ABR40251	AbR40251 Human ZAD
65	37	53.6	285	6	ABR40223	AbR40223 Human gen
66	37	53.6	285	6	ABR39852	AbR39852 ZADJ-2 po
67	37	53.6	285	6	ABR44308	AbR44308 ZADJ-2 po
68	37	53.6	285	6	ABP70912	Abp70912 ZADJ-2 pr
69	37	53.6	285	6	ABP97856	Abp97856 Amino aci
70	37	53.6	285	6	ABR40013	AbR40013 Human ZAD
71	37	53.6	285	6	ABR43908	AbR43908 ZADJ-2 po
72	37	53.6	285	6	ABP71425	Abp71425 Human GMG
73	37	53.6	285	6	ABP98487	Abp98487 Amino aci
74	37	53.6	285	6	ADA00741	Ada00741 ZADJ-2 po
75	37	53.6	285	6	ABR40144	AbR40144 ZADJ-2 pr
76	37	53.6	285	6	ABR40085	AbR40085 ZADJ-2 pr
77	37	53.6	285	6	ABR40218	AbR40218 Human cob
78	37	53.6	285	6	ABR39866	AbR39866 ZADJ-2 po
79	37	53.6	285	6	ABR40298	AbR40298 Human ZAD
80	37	53.6	285	6	AAU16581	Aau16581 Energen-r
81	37	53.6	285	6	AAU79991	Aau79991 Human ZAD
82	37	53.6	285	6	ABP71423	Abp71423 ZADJ-2 po
83	37	53.6	285	6	ABR61853	AbR61853 REDAX pol
84	37	53.6	285	6	AAO23088	Aao23088 ZADJ-2 pr
85	37	53.6	285	6	AAO23097	Aao23097 ZADJ-2 pr
86	37	53.6	285	6	AAO22230	Aao22230 Glucoset
87	37	53.6	285	6	ABR61770	AbR61770 PROLIXIN
88	37	53.6	285	6	AAO29603	Aao29603 XITAB ago
89	37	53.6	285	6	AAO29595	Aao29595 DITACIN a
90	37	53.6	285	7	ABR63587	AbR63587 ZADJ-2 pr
91	37	53.6	285	7	ABR61800	AbR61800 RYZN poly
92	37	53.6	285	7	ABR82290	AbR82290 BROMIX 11
93	37	53.6	285	7	ABR61982	AbR61982 DEKAR pol
94	37	53.6	286	6	ABR63592	AbR63592 ZADJ-2 pr
95	37	53.6	294	3	AAU76039	Aau76039 Rat skin
96	37	53.6	294	4	AAU55978	Aau55978 Skin cell
97	37	53.6	294	5	ABM72178	Abm72178 Rat prote
98	37	53.6	294	6	ABP70818	Abp70818 Murine C1

99 37 53.6 308 4 AAG89838 C glutami
100 37 53.6 314 5 ABP41487 Human ova
101 37 53.6 330 6 ABP70817 Human ClQ
102 37 53.6 330 6 ABU99123 Novel hum
103 37 53.6 331 4 AAU19557 Human dia
104 37 53.6 384 5 ABB54172 Lactococ
105 37 53.6 487 4 ABB58645 Drosophil
106 37 53.6 513 4 AAE03645 Human ext
107 37 53.6 543 4 AAB95191 Human pro
108 37 53.6 624 4 AAG08697 Novel hum
109 37 53.6 655 2 AAR31041 srnR poly
110 37 53.6 691 4 AAG05662 Novel hum
111 37 53.6 1129 4 AAM78567 Human pro
112 37 53.6 1143 4 AAG05661 Novel hum
113 37 53.6 1143 4 AAG17016 Human pro
114 37 53.6 1189 4 AAM79551 Human pro
115 37 53.6 1218 3 AAB07563 Protein e
116 37 53.6 1715 4 AAM39025 Human pol
117 37 53.6 1715 4 AAM38993 Human pol
118 37 53.6 1715 5 AAE25144 Human ARM
119 37 53.6 1715 5 AAU96840 Human kid
120 37 53.6 1715 6 AAE32128 Human cyt
121 37 53.6 1753 6 ADA09888 Human rec
122 37 53.6 1771 7 ADC06847 Human Kid
123 37 53.6 2444 4 ABB71807 Drosophil
124 36 52.2 101 5 ABP00806 Human ORF
125 36 52.2 121 3 AAB40457 Human ORF
126 36 52.2 121 5 ABP03807 Human ORF
127 36 52.2 143 2 AAR73995 Mouse WAP
128 36 52.2 165 4 AAB63835 Human pro
129 36 52.2 171 4 AAU51854 Propionib
130 36 52.2 171 6 ABM48373 Propionib
131 36 52.2 183 7 ADC86559 Human GPC
132 36 52.2 197 6 ABU41732 Protein e
133 36 52.2 247 2 AAY37186 Protein i
134 36 52.2 252 3 AAG43451 Arabidops
135 36 52.2 252 3 AAG17081 Arabidops
136 36 52.2 255 3 AAB42698 Human ORF
137 36 52.2 256 6 ADA55202 Human pro
138 36 52.2 265 6 ADA55195 Human pro
139 36 52.2 300 2 AAW73518 KSHV asse
140 36 52.2 305 2 AAW73519 KSHV asse
141 36 52.2 315 3 AAG17080 Arabidops
142 36 52.2 315 3 AAG43450 Arabidops
143 36 52.2 341 3 AAG17079 Arabidops
144 36 52.2 341 3 AAG43449 Arabidops
145 36 52.2 380 4 AAG09344 Novel hum
146 36 52.2 392 3 AAG31490 Arabidops
147 36 52.2 415 3 AAG31489 Arabidops
148 36 52.2 417 2 AAW01506 Human pan
149 36 52.2 417 2 AAW01507 Human pan
150 36 52.2 417 2 AAW01505 Wild-type

ALIGNMENTS

RESULT 1
ID AAB68087 standard; protein; 249 AA.
XX AC AAB68087;
XX AC AAB68087;
DT 09-JUL-2001 (first entry)
XX DE An anti-alpha6beta4 integrin light chain linked to a heavy chain.
XX DE Gastrointestinal epithelial tumour cell; alpha6beta4 integrin;
KW tumour-associated antigen; metastatic disease; malignant disease;
KW tumour typing; tumour screening; tumour.
XX OS Synthetic.
OS Macaca fascicularis.

Query Match 100.0%; Score 69; DR 4; Length 249;
Best Local Similarity 100.0%; Pred. No. 0.0031;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 EGGPGTTSNRLDA 13
|||
Db 226 EGGPGTTSNRLDA 238
|||
RESULT 2
AAU64492
ID AAU64492 standard; protein; 527 AA.

XX FH Location/Qualifiers
23..33
/note= "Complementarity determining region (CDR) 1 of the
light chain"
FT FT 49..55
/note= "Complementarity determining region (CDR) 2 of the
light chain"
FT FT 88..98
/note= "Complementarity determining region (CDR) 3 of the
light chain"
FT FT 110..127
/note= "linker"
FT FT 150..162
/note= "Complementarity determining region (CDR) 1 of the
heavy chain"
FT FT 177..193
/note= "Complementarity determining region (CDR) 1 of the
heavy chain"
FT FT 226..238
/note= "Complementarity determining region (CDR) 1 of the
heavy chain"
WO200130854-A2.
03-MAY-2001.
26-OCT-2000; 2000WO-SR002082.
28-OCT-1999; 99SE-00003895.
(ACTI-) ACTIVE BIOFTECH AB.
Brodin TN, Karlstroem PJ, Ohlsson LG, Tordsson MJ, Kearney PP;
Nilsson BHK;
WPI; 2001-308619/32.
N-PSDB; AAF84797.
Novel antibody for diagnosis, treatment of human metastatic and malignant
diseases, has binding structure for target structure displayed on cell
surface of human gastrointestinal epithelial tumor and normal cells.
Claim 1; Page 55-56; 75pp; English.
The present sequence represents a Monkey antibody light chain linked to a
heavy chain. The antibody binds to a target structure displayed in and on
the cell surface of human gastrointestinal epithelial tumour cells and in
a subpopulation of normal human gastrointestinal epithelial cells. The
target structure especially comprises alpha6beta4 integrin. This is a
tumour-associated antigen. The antibody, and its fragments, are useful
for treating conditions based on anti-angiogenic mechanism in humans.
They are useful for treating human metastatic and malignant disease, for
in vitro, in vivo diagnosis and prognosis of human malignant disease,
comprising tumour typing, tumour screening, tumour diagnosis and
prognosis and monitoring premalignant conditions. Quantitative in vivo
diagnosis is carried out by determining the localization of antibody to
tumour deposits in humans
Sequence 249 AA;
Query Match 100.0%; Score 69; DR 4; Length 249;
Best Local Similarity 100.0%; Pred. No. 0.0031;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX AAU64492;
AC 27-FEB-2002 (first entry)
XX
XX Propionibacterium acnes immunogenic protein #25388.
XX
XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
XX Propionibacterium acnes.
OS
XX WO200181581-A2.
XX
XX 01-NOV-2001.
XX
XX 20-APR-2001; 2001WO-US012865.
XX
XX 21-APR-2000; 2000US-0199047P.
PR
XX 02-JUN-2000; 2000US-0208841P.
PR
XX 07-JUL-2000; 2000US-0216747P.
PR
XX (CORI-) CORIXA CORP.
PA
XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
XX L'maisonneuve J, Zhang Y, Jen S, Carter D;
PI
XX WPI; 2001-616774/71.
DR
XX N-PSDB; AAS59645.
DR
XX Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.
XX
XX Example 1; SEQ ID NO 25687; 1069pp; English.
XX
XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 527 AA;
XX
XX Query Match 63.8%; Score 44; DB 4; Length 527;
XX Best Local Similarity 80.0%; Pred. No. 92;
XX Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
XX
XX 2 GGPGTTSNRL 11
XX ||||| :||
XX 7 GGPGTASHRL 16
XX
XX RESULT 3
XX ABM64701
XX ID ABM64701 standard; protein; 527 AA.

XX AC ABM64701;
XX
XX 20-OCT-2003 (first entry)
XX
XX Propionibacterium acnes immunogenic polypeptide #29377.
XX
XX Acne vulgaris; antiseborrheic; dermatological; antibacterial;
KW immunostimulant; immune response; vaccine; immunogenic.
XX
XX Propionibacterium acnes.
OS
XX WO2003033515-A1.
XX
XX 24-APR-2003.
XX
XX 11-OCT-2002; 2002WO-US032727.
XX
XX 15-OCT-2001; 2001US-00978825.
XX
XX (CORI-) CORIXA CORP.
PA
XX Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
PI Barth B, Vallie-Douglass J;
XX
XX WPI; 2003-381789/36.
DR
XX New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
PT or for stimulating an immune response specific for a P. acnes protein.
XX
XX Claim 7; SEQ ID NO 29377; 1481pp; English.
XX
XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC encoding a Propionibacterium acnes protein. The invention also relates to
CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
CC immunogenic fragments of P. acnes polypeptides. The invention
CC additionally encompasses expression vectors and host cells comprising a
CC polynucleotide of the invention; antibodies against polypeptides of the
CC invention; fusion proteins comprising a polypeptide of the invention; a
CC method for stimulating an immune response specific for a P. acnes
CC polypeptide and an isolated T cell population comprising T cells prepared
CC via this method; a vaccine composition (comprising P. acnes polypeptides,
CC polynucleotides, antibodies, fusion proteins, T cell populations, or
CC antigen-presenting cells that express the polypeptide); a method and kit
CC for detecting or determining the presence or absence of P. acnes in a
CC patient; and a method for inhibiting the development of P. acnes in a
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating acne
CC vulgaris, or for stimulating an immune response specific for a P. acnes
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridisation. The vaccine composition is useful for the
CC stimulation of an immune response against P. acnes, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a specifically claimed P. acnes polypeptide which is
CC thought to contain an immunogenic region. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 527 AA;
XX
XX Query Match 63.8%; Score 44; DB 6; Length 527;
XX Best Local Similarity 80.0%; Pred. No. 92;
XX Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
XX
XX 2 GGPGTTSNRL 11
XX ||||| :||
XX 7 GGPGTASHRL 16
XX
XX RESULT 3
XX ABM64701
XX ID ABM64701 standard; protein; 527 AA.

Db 7 GGGGTASHRL 16

RESULT 4
ABM61011
ID ABM61011 standard; protein; 527 AA.
XX
AC ABM61011;
XX
XX 20-OCT-2003 (first entry)
XX
DE Propionibacterium acnes predicted ORF-encoded polypeptide #25687.
XX
DE Acne vulgaris; antisporrheic; dermatological; antibacterial;
KW Immunostimulant; immune response; vaccine.
XX
XX Propionibacterium acnes.
OS
XX WO2003033515-A1.
PN
XX
XX 24-APR-2003.
PD
XX
XX 11-OCT-2002; 2002WO-US032727.
PF
XX
XX 15-OCT-2001; 2001US-00978825.
PR
XX
XX (CORI-). CORIXA CORP.
PA
XX Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
PI Zhang Y, Wang S, Jen S, Lodes WJ, Benson DR, Jones R, Carter D;
PI Barth B, Valliave-Douglass J;
PI
XX WPI; 2003-381789/36.
DR N-PSDB; ACF64574.
DR
XX
XX New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
PT or for stimulating an immune response specific for a P. acnes protein.
XX
XX Example 1; SEQ ID NO 25687; 1481pp; English.
XX
XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC encoding a Propionibacterium acnes protein. The invention also relates to
CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
CC immunogenic fragments of P. acnes polypeptides. The invention
CC additionally encompasses expression vectors and host cells comprising a
CC polynucleotide of the invention; antibodies against polypeptides of the
CC invention; fusion proteins comprising a polypeptide of the invention; a
CC method for stimulating an immune response specific for a P. acnes
CC polypeptide and an isolated T cell population comprising T cells prepared
CC via this method; a vaccine composition (comprising P. acnes polypeptides,
CC polynucleotides, antibodies, fusion proteins, T cell populations, or
CC antigen-presenting cells that express the polypeptide); a method and kit
CC for detecting or determining the presence or absence of P. acnes in a
CC patient; and a method for inhibiting the development of P. acnes in a
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating acne
CC vulgaris, or for stimulating an immune response specific for a P. acnes
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridisation. The vaccine composition is useful for the
CC stimulation of an immune response against P. acnes, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a polypeptide predicted to be encoded by an ORF (open
CC reading frame) contained within the P. acnes polynucleotides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

Query Match 63.8%; Score 44; DB 6; Length 527;
Best Local Similarity 80.0%; Pred. No. 92;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

2 GGGGTASHRL 11

Db 7 GGGGTASHRL 16

RESULT 5
ABM60469
ID ABB60469 standard; protein; 749 AA.
XX
AC ABB60469;
XX
XX 26-MAR-2002 (first entry)
DT
XX Drosophila melanogaster polypeptide SEQ ID NO 8199.
DE
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
XX Drosophila melanogaster.
OS
XX WO200171042-A2.
PN
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US009231.
PF
XX
XX 23-MAR-2000; 2000US-0191637P.
PR
XX 11-JUL-2000; 2000US-00614150.
PR
XX (PEKE) PE CORP NY.
PA
XX Venter JC, Adams M, Li PWD, Myers EW;
PI
XX WPI; 2001-656860/75.
DR N-PSDB; ABL04572.
DR
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
XX Disclosure; SEQ ID NO 8199; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

Query Match 62.3%; Score 43; DB 4; Length 749;
Best Local Similarity 66.7%; Pred. No. 1.9e+02;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

2 GGGGTASHRLDA 13

Db 127 GGGGTASHRLDA 138

RESULT 6
ABP76682
ID ABP76682 standard; protein; 19938 AA.
XX
AC ABP76682;
XX
XX 26-FEB-2003 (first entry)
DT
XX Streptomyces viridochromogenes Avi gene cluster polypeptide frame 6.
DE
XX

KW Avilamycin; antibacterial; virucide; protozoacide; fungicide; infection;
 KW medicine; Staphylococcus aureus; biosynthetic gene cluster; enzyme.
 XX
 OS Streptomyces viridochromogenes.
 XX
 XX WO200268436-A1.
 XX

XX 06-SEP-2002.
 XX
 XX 24-AUG-2001; 2001WO-EP009815.
 XX
 XX 25-FEB-2001; 2001DE-01009166.
 XX
 XX (COMB-) COMBINATURE BIOPHARM AG.
 XX

XX Weitnauer G, Muehlenweg A, Trefzer A, Bechtold A;
 XX WPI; 2003-018650/01.
 XX N-PSDB; ABZ37516.
 XX

XX New avilamycin derivatives, useful for treatment of infections, and
 XX nucleic acid encoding avilamycin synthesis enzymes.
 XX

XX Example 1; Page 68-301; 319pp; German.
 XX
 XX The invention relates to avilamycin derivatives (I) with antibacterial,
 XX virucide, protozoacide and fungicide activity. (I) are useful for
 XX treatment of infections (bacterial, viral, protozoal or fungal), in human
 XX or veterinary medicine, particularly where caused by Staphylococcus
 XX aureus. (I) are more hydrophilic than known avilamycins. The present
 XX sequence is that of an avilamycin synthesis enzyme from the Streptomyces
 XX viridochromogenes Avilamycin A biosynthetic gene cluster (ABZ37515-
 XX ABZ37516)
 XX

SQ Sequence 19938 AA;

Query Match 59.4%; Score 41; DB 6; Length 19938;
 Best Local Similarity 66.7%; Pred. No. 1.2e+04;
 Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 GGGGTTSNRLDA 13
 |||||
 DB 4312 GGGGTAEPFRDA 4323

RESULT 7
 AAY74086
 ID AAY74086 standard; protein; 224 AA.
 XX
 XX AAY74086;
 XX
 XX 14-MAR-2000 (first entry)
 XX
 XX Human prostate tumor EST fragment derived protein #73.
 DE
 XX Pancreas; tumor; EST; expressed sequence tag; human; cytostatic;
 KW treatment.
 KW
 XX Homo sapiens.
 OS
 XX DE19820190-A1.
 XX
 XX 04-NOV-1999.
 XX
 XX 28-APR-1998; 98DE-01020190.
 XX
 XX 28-APR-1998; 98DE-01020190.
 XX
 XX (META-) METAGEN GES GENOMFORSCHUNG MBH.
 PA
 XX Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl B;
 XI
 XX WPI; 1999-621386/54.
 XX

DR N-PSDB; AAZ52948.
 XX
 XX New human nucleic acid sequences from pancreatic tumors, and related
 PT proteins.
 PT
 XX Claim 23; Page 422; 502pp; German.
 XX

XX This invention describes novel polypeptides and their encoding nucleic
 CC acids derived from human pancreatic tumor tissue which have cytostatic
 CC activity. The sequences are also useful in producing pharmaceutical
 CC compositions for treatment of pancreatic tumors. AAY73814-Y74252
 CC represent protein fragments encoded by the human pancreatic tumor cDNA
 CC library derived expressed sequence tag (EST) sequences represented in
 CC AAZ52858-Z53014
 XX

SQ Sequence 224 AA;

Query Match 58.0%; Score 40; DB 2; Length 224;
 Best Local Similarity 77.8%; Pred. No. 1.7e+02;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EGGPGTTSN 9
 |||||
 DB 91 EGGPGTTSN 99

RESULT 8
 AAM85724
 ID AAM85724 standard; protein; 77 AA.
 XX

XX AAM85724;
 XX
 XX 07-NOV-2001 (first entry)
 DT
 XX Human immune/haematopoietic antigen SEQ ID NO:13317.
 DE
 XX

XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
 KW cytostatic; gene therapy; vaccine; metastasis.
 XX
 XX Homo sapiens.
 OS

XX WO200157182-A2.
 PN
 XX 09-AUG-2001.
 PD
 XX

XX 17-JAN-2001; 2001WO-US001354.
 XX
 XX 31-JAN-2000; 2000US-0179065P.
 PR
 XX 04-FEB-2000; 2000US-0180628P.
 PR
 XX 24-FEB-2000; 2000US-0184664P.
 PR
 XX 02-MAR-2000; 2000US-0186350P.
 PR
 XX 16-MAR-2000; 2000US-0189874P.
 PR
 XX 17-MAR-2000; 2000US-0190076P.
 PR
 XX 18-APR-2000; 2000US-0198123P.
 PR
 XX 19-MAY-2000; 2000US-0205515P.
 PR
 XX 07-JUN-2000; 2000US-0209467P.
 PR
 XX 28-JUN-2000; 2000US-0214886P.
 PR
 XX 30-JUN-2000; 2000US-0215135P.
 PR
 XX 07-JUL-2000; 2000US-0216647P.
 PR
 XX 07-JUL-2000; 2000US-0216880P.
 PR
 XX 11-JUL-2000; 2000US-0217487P.
 PR
 XX 11-JUL-2000; 2000US-0217496P.
 PR
 XX 14-JUL-2000; 2000US-0218290P.
 PR
 XX 26-JUL-2000; 2000US-0220963P.
 PR
 XX 26-JUL-2000; 2000US-0220964P.
 PR
 XX 14-AUG-2000; 2000US-0224518P.
 PR
 XX 14-AUG-2000; 2000US-0224519P.
 PR
 XX 14-AUG-2000; 2000US-0225213P.
 PR
 XX 14-AUG-2000; 2000US-0225214P.
 PR
 XX 14-AUG-2000; 2000US-0225266P.
 PR
 XX 14-AUG-2000; 2000US-0225267P.
 PR
 XX 14-AUG-2000; 2000US-0225268P.
 PR
 XX 14-AUG-2000; 2000US-0225270P.
 PR

PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 14-AUG-2000; 2000US-0226279P.
PR 14-AUG-2000; 2000US-0226281P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 23-AUG-2000; 2000US-0228924P.
PR 30-AUG-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232337P.
PR 14-SEP-2000; 2000US-0232338P.
PR 14-SEP-2000; 2000US-0232339P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 21-SEP-2000; 2000US-0234977P.
PR 25-SEP-2000; 2000US-0234988P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.

PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 03-DEC-2000; 2000US-0251030P.
PR 03-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-483426/52.
XX N-PSDB; AAK58505.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
XX Claim 11; SEQ ID NO 13317; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
XX amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
XX activity, and can be used in gene therapy and vaccine production. (I)
XX proteins and polynucleotides may be used in the prevention, diagnosis and
XX treatment of diseases associated with inappropriate (I) expression. For
XX example, they may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of (I) by expressing inactive proteins or to
XX supplement the patients own production of (I). Additionally, (I)
XX nucleic acids may be used to produce the secreted (I), by inserting the
XX protein. (I) proteins and polynucleotides may be used to prevent,
XX cancers and cancer metastases of hematopoietic-related diseases, especially
XX to AAK87694 represent human immune/hematopoietic antigen genomic
XX sequences from the present invention. AAK54942 to AAK54950 and AAK82169
XX represent sequences used in the exemplification of the present invention
XX
XX SQ Sequence 77 AA;

Query Match 56.5%; Score 39; DB 4; Length 77;
Best Local Similarity 63.6%; Pred. No. 84;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EGGPGTTSNRL 11
| | | | | | | | | |
Db 47 EGGPGRASRRV 57

RESULT 9
ABU41557
ID ABU41557 standard; protein; 130 AA.
XX
XX
AC ABU41557;
XX
XX 19-JUN-2003 (first entry)
XX
XX Protein encoded by prokaryotic essential gene #27084.
XX
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
XX Pseudomonas syringae.
XX
XX WO200277183-A2.
XX
XX 03-OCT-2002.
XX
XX 21-MAR-2002; 2002WO-US009107.
XX
XX 21-MAR-2001; 2001US-00815242.
XX
XX 06-SEP-2001; 2001US-00948993.
XX
XX 25-OCT-2001; 2001US-0342923P.
XX
XX 08-FEB-2002; 2002US-00072851.
XX
XX 06-MAR-2002; 2002US-0362699P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI: 2003-029926/02.
XX
XX N-PSDB; ACA45427.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
XX for homologous nucleic acids required for cellular proliferation to
XX isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 25; SEQ ID NO 69481; 1766pp; English.

CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of

CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 130 AA;
QY Query Match 56.5%; Score 39; DB 6; Length 130;
Best Local Similarity 61.5%; Pred. No. 1.4e+02;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
Db 1 EGGPGTTSNRLDA 13
| | | | | | | | | |
11 EGTGSGTPRELGA 23

RESULT 10
AAU83139
ID AAU83139 standard; protein; 219 AA.
XX
XX AAU83139;
XX
XX 08-MAY-2002 (first entry)
XX
XX Novel secreted protein Z838027G3P.
XX
XX Protein secretion; mammalian secreted polypeptide; MSP.
XX
XX Homo sapiens.
XX
XX WO200202621-A2.
XX
XX 10-JAN-2002.
XX
XX 28-JUN-2001; 2001WO-US020638.
XX
XX 30-JUN-2000; 2000US-0215446P.
XX
XX (ZYMO) ZYMOGENETICS INC.
XX
XX Sheppard PO, Presnell SR;
XX WPI: 2002-147999/19.
XX
XX N-PSDB; ABK33054.
XX
XX Novel isolated mammalian secreted polypeptide useful in therapeutic and
XX diagnostic methods, to direct secretion of other proteins of interest
XX from host cell, as educational tools, and as laboratory practicum kits.
XX
XX Claim 12; Page 194-195; 397pp; English.

CC The invention describes an isolated mammalian secreted polypeptide (MSP)
CC (1). (1) is useful to direct the secretion of other proteins of interest
CC from a host cell, to monitor secretion of proteins, to degenerate
CC sequences comprising all nucleotide sequences encoding a particular
CC polypeptide, to screen for cell metabolism effecting receptors, for
CC identifying new target receptors and drug design, for identifying, for
CC protein purification, for determining the weight of expressed MSP
CC polypeptides as a ratio to total protein expressed, for identifying
CC peptide cleavage sites, for coupling amino and carboxy terminal tags, for
CC amino acid sequence analysis, for monitoring biological activities of the
CC protein in vitro and in vivo, and to teach analytical skills and as
CC reagents for the study of cells, receptors, and other binding molecules.
CC The polynucleotide is useful for radiation hybrid mapping, and somatic
CC cell genetic technique developed for constructing high-resolution,
CC contiguous maps of mammalian chromosomes. Reagents disclosed in the
CC invention may be used to detect metabolic abnormalities characterised by
CC over or under production of the protein. This is the amino acid sequence
CC of a mammalian secreted polypeptide, described in the method of the
XX invention
XX Sequence 219 AA;

Query Match 56.5%; Score 39; DB 5; Length 219;
 Best Local Similarity 46.2%; Pred. No. 2.5e+02;
 Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 EGGPGTTSNRDLA 13
 :|||||:|:
 Db 33 QGGFGSTNSKRQA 45

RESULT 11
 ID ABP43778 standard; protein; 296 AA.
 XX
 AC ABP43778;
 XX
 DT 26-FEB-2003 (first entry)
 XX
 DE FLJ14840 fis clone.
 XX
 KW Neuroprotective; immunomodulator; cancer; chromosome 15; cytostatic;
 KW anti-inflammatory; gene therapy; nutritional supplement; wound; burn;
 KW ulcer; Alzheimer's disease; Huntington's disease;
 KW amyotrophic lateral sclerosis; autoimmune disorder; inflammation;
 KW vulnerary.
 XX
 OS Homo sapiens.
 XX
 PN WO200231111-A2.
 XX
 PD 18-APR-2002.
 XX
 PF 11-OCT-2001; 2001WO-US027760.
 XX
 PR 12-OCT-2000; 2000US-00687527.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
 PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;
 XX
 DR WPI; 2002-426278/45.
 DR N-PSDB; ABQ61022.
 XX
 PT New polypeptides and their encoded proteins, useful as nutritional
 PT sources or supplements, or in gene therapy, particularly for treating
 PT wounds, Alzheimer's disease, amyotrophic lateral sclerosis, cancer or
 PT inflammation.
 XX
 PS Claim 20; SEQ ID # 681; 357pp + Sequence Listing; English.
 XX
 CC The invention relates to 446 newly isolated polynucleotide sequences. The
 CC activity of polynucleotides of the invention may be described as,
 CC vulnerary, neuroprotective, immunomodulator, cytostatic and anti-
 CC inflammatory. Compositions comprising nucleic acids of the invention are
 CC useful for treating a mammalian subject, or as nutritional sources or
 CC supplements. These are useful in gene therapy, particularly for treating
 CC wounds, burns or ulcers, Alzheimer's disease, Huntington's disease,
 CC amyotrophic lateral sclerosis, autoimmune disorders, cancer or
 CC inflammation. The nucleic acids and polypeptides are also useful in
 CC diagnostic and research methods. The sequences given in records ABP43544-
 CC ABP43989 represent polypeptides encoded by polynucleotides of the
 CC invention. NOTE: the sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 296 AA;

Query Match 56.5%; Score 39; DB 5; Length 296;
 Best Local Similarity 46.2%; Pred. No. 3.4e+02;
 Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 EGGPGTTSNRDLA 13
 :|||||:|:
 Db 33 QGGFGSTNSKRQA 45

RESULT 12
 ID ABB68531 standard; protein; 302 AA.
 XX
 AC ABB68531;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 32385.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US009231.
 XX
 PR 23-MAR-2000; 2000US-0191637P.
 XX
 PR 11-JUL-2000; 2000US-00614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 PI WPI; 2001-656860/75.
 DR N-PSDB; ABL12634.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions.
 XX
 PS Disclosure; SEQ ID NO 32385; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 302 AA;

Query Match 56.5%; Score 39; DB 4; Length 302;
 Best Local Similarity 54.5%; Pred. No. 3.5e+02;
 Matches 6; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 3 GPGTTSNRDLA 13
 :|||||:|:
 Db 71 GPGSSSSKLDS 81

RESULT 13
 ID AAB94593 standard; protein; 396 AA.
 XX
 AC AAB94593;
 XX
 DT 26-JUN-2001 (first entry)
 XX
 DE Human protein sequence SEQ ID NO:15405.
 XX
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.

```

XX OS Homo sapiens.
XX PN EP1074617-A2.
XX PD 07-FEB-2001.
XX PF 28-JUL-2000; 2000EP-00116126.
XX PR 29-JUL-1999; 99JP-00248036.
XX PR 27-AUG-1999; 99JP-00300253.
XX PR 11-JAN-2000; 2000JP-00118776.
XX PR 02-MAY-2000; 2000JP-00183767.
XX PR 09-JUN-2000; 2000JP-00241899.
XX PA (HELI-) HELIX RES INST.
XX PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX DR WPI; 2001-318749/34.
XX PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
XX PT length cDNAs defined in the specification, and for the detection and/or
XX PT diagnosis of the abnormality of the proteins encoded by the full-length
XX PT cDNAs.
XX PS Claim 8; SEQ ID NO 15405; 2537pp + Sequence Listing; English.
XX CC The present invention describes primer sets for synthesizing 5602 full-
XX CC length cDNAs defined in the specification. Where a primer set comprises:
XX CC (a) an oligo-dr primer and an oligonucleotide complementary to the
XX CC complementary strand of a polynucleotide which comprises one of the 5602
XX CC nucleotide sequences defined in the specification, where the
XX CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX CC of an oligonucleotide comprising a sequence complementary to the
XX CC complementary strand of a polynucleotide which comprises a 5'-end
XX CC sequence and an oligonucleotide comprising a sequence complementary to a
XX CC polynucleotide which comprises a 3'-end sequence, where the
XX CC oligonucleotide comprises at least 15 nucleotides and the combination of
XX CC the 5'-end sequence/3'-end sequence is selected from those defined in the
XX CC specification. The primer sets can be used in antisense therapy and in
XX CC gene therapy. The primers are useful for synthesizing polynucleotides,
XX CC particularly full-length cDNAs. The primers are also useful for the
XX CC detection and/or diagnosis of the abnormality of the proteins encoded by
XX CC the full-length cDNAs. The primers allow obtaining of the full-length
XX CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
XX CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893
XX CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
XX CC oligonucleotides, all of which are used in the exemplification of the
XX CC present invention
XX SQ Sequence 396 AA;
    Query Match 56.5%; Score 39; DB 4; Length 396;
    Best Local Similarity 46.2%; Pred. No. 4.6e+02;
    Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 EGGPGTTSNRLDA 13
    :||||:|:::|
DB . 210 QGGPGSTNSKROA 222

RESULT 14
AAG67257
ID AAG67257 standard; protein; 396 AA.
XX AC AAG67257;
XX AC AAG67257;
XX DT 13-NOV-2001 (first entry)
XX DE Amino acid sequence of a human liver-associated gene.

QY 1 EGGPGTTSNRLDA 13
    :||||:|:::|
DB . 210 QGGPGSTNSKROA 222

RESULT 15
AAU35107
ID AAU35107 standard; protein; 454 AA.
XX AC AAU35107;
XX AC AAU35107;
XX DT 13-FEB-2002 (first entry)
XX DE Enterococcus faecalis cellular proliferation protein #394.
XX DE Antisense; prokaryotic cellular proliferation protein; antibiotic;
XX KW antibacterial; drug design.
XX OS Enterococcus faecalis.
XX PN WO200170955-A2.
XX PD 27-SEP-2001.
XX PF 21-MAR-2001; 2001WO-US009180.
XX PR 21-MAR-2000; 2000US-0191078P.

```

Liver-associated gene; hepatic cancer; liver cancer; marker.

Homo sapiens.

WO200109318-A1.

08-FEB-2001.

28-JUL-2000; 2000WO-JP005064.

29-JUL-1999; 99JP-00248036.

27-AUG-1999; 99JP-00300253.

18-OCT-1999; 99US-0159590P.

11-JAN-2000; 2000JP-00118776.

17-FEB-2000; 2000US-0183322P.

02-MAY-2000; 2000JP-00183767.

09-JUN-2000; 2000JP-00241899.

(HELI-) HELIX RES INST.

Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T, Aburatani H;
Kodama T, Hippi Y, Taniguchi H;

WPI; 2001-541221/60.

N-PSDB; AAH77810.

New genes and their proteins found in hepatic cancer, vectors,
antibodies, and screening method for compounds with binding activity,
useful as diagnostics and targets for treatment agents.

Claim 1; Page 79-80; 99pp; Japanese.

The present sequence is encoded by a liver-associated gene. The
expression level of this gene changes in liver (hepatic) cancer. Liver-
associated gene can be used as markers in blood tests for screening for
the early stages of the liver cancer. The encoded proteins and peptides
can also be used as targets for screening for compounds to treat the
disease. They can also be used for predicting the progress of the disease

Sequence 396 AA;

Query Match 56.5%; Score 39; DB 4; Length 396;
Best Local Similarity 46.2%; Pred. No. 4.6e+02;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 EGGPGTTSNRLDA 13
:||||:|:::|

DB 210 QGGPGSTNSKROA 222

RESULT 15
AAU35107
ID AAU35107 standard; protein; 454 AA.

AC AAU35107;
AC AAU35107;

DT 13-FEB-2002 (first entry)

DE Enterococcus faecalis cellular proliferation protein #394.
Antisense; prokaryotic cellular proliferation protein; antibiotic;
antibacterial; drug design.

OS Enterococcus faecalis.

PN WO200170955-A2.

PD 27-SEP-2001.

PF 21-MAR-2001; 2001WO-US009180.

PR 21-MAR-2000; 2000US-0191078P.

```

PR 23-MAY-2000; 2000US-0206848P.
PR 26-MAY-2000; 2000US-0207272P.
PR 23-OCT-2000; 2000US-0242578P.
PR 27-NOV-2000; 2000US-0253625P.
PR 22-DEC-2000; 2000US-0257931P.
PR 16-FEB-2001; 2001US-0269308P.
XX (ELIT-) ELITRA PHARM INC.
XX
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX WPI; 2001-611495/70.
DR N-PSDB; AAS52966.
XX
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids.
XX
XX Example 3; SEQ ID NO 10700; 511pp; English.
XX
XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the genes,
CC their use in the discovery of novel antibiotics, the essential genes
CC themselves and the encoded proteins. The prokaryotes used are Escherichia
CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence represents an essential prokaryotic
CC cellular proliferation protein. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 454 AA;
SQ
Query Match 56.5%; Score 39; DB 4; Length 454;
Best Local Similarity 80.0%; Pred. No. 5.3e+02;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 GGGGTTSNRL 11
DB 152 GRPGTFSNRL 161
RESULT 16
ABU29209
ID ABU29209 standard; protein; 454 AA.
XX
AC ABU29209;
XX
DT 19-JUN-2003 (first entry)
DE
DE Protein encoded by Prokaryotic essential gene #14736.
XX
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Enterococcus faecalis.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
XX 21-MAR-2002; 2002WO-US009107.
XX
XX 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX (ELIT-) ELITRA PHARM INC.
XX
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
DR N-PSDB; ACA33079.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 25; SEQ ID NO 57133; 1766pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway; (8)
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 454 AA;
SQ
Query Match 56.5%; Score 39; DB 6; Length 454;
Best Local Similarity 80.0%; Pred. No. 5.3e+02;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 GGGGTTSNRL 11
DB 152 GRPGTFSNRL 161
RESULT 17
AAB85503
ID AAB85503 standard; protein; 600 AA.
XX
XX AAB85503;
XX
DT 25-SEP-2001 (first entry)
XX
DE Human protein kinase SGK146.
XX
XX Protein kinase; enzyme; cytostatic; neurotropic; neuroprotective; human;
KW antiparkinsonian; virucide; antibacterial; antifungal; antimigraine;
KW analgesic; hypotensive; hypertensive; immunosuppressive; antiallergic;

```

KW antipsoriatic; antirheumatic; antiarthritic; ophthalmological; anorectic;
 KW osteopathic; thrombolytic; antiarteriosclerotic; antiasthmatic;
 KW vasotropic; antidiabetic; gene therapy.
 OS Homo sapiens.
 XX WO200115356-A2.
 PN 02-AUG-2001.
 XX 25-JAN-2001; 2001WO-US002337.
 XX 25-JAN-2000; 2000US-0178078P.
 PR 31-JAN-2000; 2000US-0179364P.
 PR 17-FEB-2000; 2000US-0183173P.
 PR 17-MAR-2000; 2000US-0190162P.
 PR 29-MAR-2000; 2000US-0193404P.
 PR 13-NOV-2000; 2000US-0247013P.
 XX (SUGEN-) SUGEN INC.
 XX Plowman G, Whyte D, Manning G, Sudarsanam S, Martinez R;
 XX WPI; 2001-476202/51.
 DR N-PSDB; AAH46903.
 XX Kinase polypeptides useful for treating cancers, Alzheimer's disease,
 PT viral infections, diabetes, obesity, organ transplant rejection and
 PT rheumatoid arthritis.
 XX Claim 7; Page 215; 218pp; English.
 PS The invention provides human protein kinases and protein kinase-like
 XX enzymes and polynucleotides encoding the polypeptides. The kinase
 CC polypeptides and their modulators are useful for treating a disease or
 CC disorder such as cancer, immune-related diseases, cardiovascular disease,
 CC brain or neuronal-associated disease and metabolic disorders, including
 CC cancers of tissues, cancers of hematopoietic origin, diseases of the
 CC central nervous system, diseases of the peripheral nervous system,
 CC Alzheimer's disease, Parkinson's disease, multiple sclerosis, amyotrophic
 CC lateral sclerosis, viral infections, infections caused by prions,
 CC bacteria and fungi, ocular diseases, migraines, pain, sexual dysfunction,
 CC mood disorders, attention disorders, cognition disorders, hypotension,
 CC hypertension, psychotic disorders, neurological disorders, dyskinesias,
 CC metabolic disorders, and organ transplant rejection. They are also useful
 CC for treating rhinitis, autoimmunity, atherosclerosis, psoriasis,
 CC osteoarthritis, asthma, chronic inflammatory pelvic disease, chronic
 CC inflammatory bowel disease, rheumatoid arthritis, metabolic disorders
 CC such as diabetes, obesity, cardiovascular diseases such as reperfusion
 CC injury, coronary thrombosis, clotting disorders and atherosclerosis,
 CC ocular diseases such as glaucoma, retinopathy and macular degeneration,
 CC psychiatric and neurological disorders such as anxiety, schizophrenia,
 CC dementia, manic depression, etc. The polynucleotides are useful in gene
 CC therapy techniques to treat the above mentioned disorders. Sequences
 CC AAB85491-85522 represent the human protein kinases of the invention
 XX Sequence 600 AA;
 SQ
 Query Match 56.5%; Score 39; DB 4; Length 600;
 Best Local Similarity 87.5%; Pred. No. 7.1e+02;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 EGGPGTTS 8
 DB 187 EGGPGGATS 194
 RESULT 18
 AAE24147
 ID AAE24147 standard; protein; 600 AA.
 XX
 AC AAE24147;
 XX

DT 23-SEP-2002 (first entry)
 DE Human kinase (PKIN)-18 protein.
 XX
 KW Human; kinase; PKIN; cancer; immune system disorder; atherosclerosis;
 KW acquired immune deficiency syndrome; AIDS; Addison's disease; allergy;
 KW asthma; multiple sclerosis; psoriasis; arteriosclerosis; cirrhosis;
 KW development; hepatitis; cardiovascular; hypertension; drug screening;
 KW myocardial infarction; Goodpasture's syndrome; lipid disorder; growth;
 KW fatty liver; Gaucher's disease; Niemann-Pick's disease; anorectic;
 KW hypercholesterolaemia; obesity; gene therapy; cytostatic; anti-HIV;
 KW neuroprotective; hepatotropic; hypotensive; cardiac; nephrotropic;
 KW hyperlipidaemia; enzyme.
 XX Homo sapiens.
 FH Key Location/Qualifiers
 FT Domain 278..535 /note= "Eukaryotic protein kinase domain"
 FT Domain 280..525 /note= "Protein kinase domain"
 FT Domain 284..526 /note= "Protein kinase domain"
 XX WO200233099-A2.
 XX 25-APR-2002.
 XX 20-OCT-2001; 2001WO-US047728.
 XX 20-OCT-2000; 2000US-0242410P.
 PR 27-OCT-2000; 2000US-0244068P.
 PR 03-NOV-2000; 2000US-0245708P.
 PR 09-NOV-2000; 2000US-0247672P.
 PR 16-NOV-2000; 2000US-0249565P.
 PR 22-NOV-2000; 2000US-0252730P.
 PR 01-DEC-2000; 2000US-0250807P.
 XX (INCV-) INCVTE GENOMICS INC.
 XX Gururajan R, Baughn MR, Wallia NK, Elliott VS, Xu Y, Arvizu C;
 PI Yao MG, Ramkumar J, Dang L, Tang Y, Hafalia AJA, Nguyen DB;
 PI Gandhi AR, Lu Y, Yue H, Burford N, Bandman O, Tribouley CM, Lal PG;
 PI Recipon SA, Lu DAM, Borowsky ML, Thornton M, Swarnaker A;
 PI Thangavelu K, Khan FA, Ison CH;
 XX WPI; 2002-454603/48.
 DR N-PSDB; AAD38861.
 XX New human kinase polypeptide, for diagnosing, preventing and treating
 PT cancer, immune system disorders, growth and development disorders,
 PT cardiovascular disorders and lipid disorders.
 XX Claim 1; Page 171-172; 210pp; English.
 XX The invention relates human kinases (PKIN) and their corresponding
 CC nucleic acid sequences. PKIN and its DNA are useful for diagnosing,
 CC treating and preventing cancer, an immune system disorder (e.g., acquired
 CC immune deficiency syndrome (AIDS), Addison's disease, allergy, asthma,
 CC atherosclerosis, multiple sclerosis, psoriasis), disorders affecting
 CC growth and development (e.g., arteriosclerosis, cirrhosis, hepatitis),
 CC cardiovascular disorder (e.g., hypertension, myocardial infarction,
 CC Goodpasture's syndrome), and a lipid disorder (e.g., fatty liver,
 CC Gaucher's disease, Niemann-Pick's disease, hypercholesterolaemia,
 CC hyperlipidaemia, obesity), and for assessing the effects of exogenous
 CC compounds. Anti-PKIN antibody is useful in a diagnostic test for a
 CC condition or a disease associated with the expression of PKIN in a
 CC biological sample. A composition comprising PKIN or an agonist or
 CC antagonist of PKIN is useful for treating a disease or condition
 CC associated with decreased or increased expression of functional PKIN.
 CC PKIN is useful in a number of drug screening techniques and to analyse
 CC the proteome of a tissue or cell type. PKIN DNA is useful for creating
 CC knockin humanised animals or transgenic animals to model human diseases,

CC and in somatic or germline gene therapy. The present sequence is human
CC PKIN protein
XX
SQ Sequence 600 AA;

Query Match 56.5%; Score 39; DB 5; Length 600;
Best Local Similarity 87.5%; Pred. No. 7.1e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EGGPGTTS 8
Db 187 EGGPGATS 194
|||||

RESULT 19
AAB58459
ID AAB58459 standard; protein; 609 AA.
XX
AC AAB58459;
XX
DT 14-MAR-2001 (first entry)
XX
DE Lung cancer associated polypeptide sequence SEQ ID 797.
XX
KW Human; lung cancer associated protein; neuroprotective; cytostatic;
KW cardioactive; immunomodulatory; muscular active; vulneryary;
KW gastrointestinal; nephrotropic; antiinfective; gynecological;
KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;
KW proliferative disorder; wound healing; infectious disease.
XX
OS Homo sapiens.
XX
XX WO200055180-A2.
PN
XX
PD 21-SEP-2000.
XX
XX
PF 08-MAR-2000; 2000WO-US005918.
XX
PR 12-MAR-1999; 99US-0124270P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PA (ROSE/) ROSEN C A.
PI Ruben SM;
XX
XX WPI; 2000-587514/55.
DR N-PSDB; AAF18335.
XX
XX Lung cancer associated gene sequences, referred to as lung cancer
PT antigens, useful for treatment, prevention, and diagnosis of disorders
PT such as lung cancer.
XX
PS Claim 11; Page 1334-1336; 1425pp; English.
XX
XX Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer
CC associated proteins represented in AAB58106 - AAB58548. Lung cancer
CC associated proteins and polynucleotide sequences, their agonists, and
CC antagonists may have neuroprotective; cytostatic; cardioactive;
CC immunomodulatory; muscular active general; vulneryary; gastrointestinal
CC general; nephrotropic; antiinfective; gynecological; or antibacterial
CC activity. The invention also includes antibodies specific for the protein
CC or polynucleotide sequences. The lung cancer associated polynucleotide
CC sequences may be used for detection of lung cancer, chromosome
CC identification, as chromosome markers, and for numerous other diagnostic
CC or research purposes. The proteins may be used to treat disorders such as
CC neural, immune, muscular, reproductive, gastrointestinal, pulmonary,
CC cardiovascular, renal, and proliferative disorders. The proteins may also
CC be used in the treatment of wounds and infectious diseases.
CC Polynucleotide sequences AAF18425 - AAF18433 and peptide AAB58549 are
CC used in the course of the invention for the identification and
CC characterisation of the polynucleotide and protein sequences
XX
SQ Sequence 609 AA;

Query Match 56.5%; Score 39; DB 3; Length 609;
Best Local Similarity 70.0%; Pred. No. 7.2e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GGPGETTSNRL 11
Db 586 GGPGGGNRL 595
|||||

RESULT 20
ABB63351
ID ABB63351 standard; protein; 1157 AA.
XX
AC ABB63351;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 16845.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
XX WO200171042-A2.
PN
XX
PD 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US009231.
XX
XX 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
PI
XX
XX WPI; 2001-656860/75.
DR N-PSDB; ABL07454.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions.
XX
XX Disclosure; SEQ ID NO 16845; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (AAB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1157 AA;

Query Match 56.5%; Score 39; DB 4; Length 1157;
Best Local Similarity 77.8%; Pred. No. 1.4e+03;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GGPGETTSNR 10
Db 772 GGPGETTQR 780
|||||

RESULT 21
ABB60494
ID ABB60494 standard; protein; 1253 AA.
XX

AC ABB60494;
 XX
 XX 26-MAR-2002 (first entry)
 XX
 XX Drosophila melanogaster polypeptide SEQ ID NO 8274.
 XX
 XX Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 KW
 XX Drosophila melanogaster.
 XX
 XX WO200171042-A2.
 PN
 XX 27-SEP-2001.
 PD
 XX
 XX 23-MAR-2001; 2001WO-US009231.
 XX
 XX 23-MAR-2000; 2000US-0191637P.
 PR
 XX 11-JUL-2000; 2000US-00614150.
 PR
 XX (PEKE) PE CORP NY.
 PA
 XX Venter JC, Adams M, Li PWD, Myers EW;
 PI
 XX WPI; 2001-656860/75.
 DR
 XX N-PSDB; ABL04597.
 DR
 XX
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signaling and cell-cell
 PT interactions.
 PT
 XX Disclosure; SEQ ID NO 8274; 21pp + Sequence Listing; English.
 PS
 XX The invention relates to an isolated nucleic acid detection reagent
 XX capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS57737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 1253 AA;
 SQ
 Query Match 56.5%; Score 39; DB 4; Length 1253;
 Best Local Similarity 66.7%; Pred. No. 1.5e+03;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 1 EGGPGTTSNRDL 12
 DB 1233 EGGPGTTSNRDL 1244
 RESULT 22
 AAB59078
 ID AAB59078 standard; protein; 102 AA.
 XX
 AC AAB59078;
 AC
 XX 27-MAR-2001 (first entry)
 DT
 XX Breast and ovarian cancer associated antigen protein sequence SEQ ID 786.
 DE
 XX Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;
 KW neurotropic; neuroprotective; antiviral; antiallergic; hepatotropic;
 KW antibacterial; antifungal; antiparasitic; cardiant; immune disorder;
 KW Addison's disease; allergy; autoimmune haemolytic anaemia;
 KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
 KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
 KW cardiovascular disorder; wound healing; neurological disease.

XX Homo sapiens.
 OS
 XX WO200055173-A1.
 PN
 XX 21-SEP-2000.
 PD
 XX 08-MAR-2000; 2000WO-US005881.
 PF
 XX 12-MAR-1999; 99US-0124270P.
 PR
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Ruben SM;
 PI
 XX WPI; 2000-611515/58.
 DR
 XX N-PSDB; AAF21981.
 DR
 XX New human breast and ovarian cancer associated gene sequences and the
 PT polypeptides encoded by these genes, useful in the prevention, treatment
 PT and diagnosis of cancer, immune disorders, cardiovascular disorders and
 PT neurological diseases.
 PT
 XX Claim 11; Page 1249-1250; 1299pp; English.
 PS
 XX Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
 CC proteins AAB59711 - AAB59128. The DNA and protein sequences are
 CC associated with breast and ovarian cancer. Included in the invention are
 CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the
 CC isolation and characterisation of the DNA and protein sequences of the
 CC invention. The breast and ovarian cancer associated DNA, protein, agonist
 CC or antagonist sequences exhibit cytostatic; immunosuppressive; neurotropic;
 CC neuroprotective; antiviral; antiallergic; hepatotropic; antidiabetic;
 CC antiinflammatory; antiulcer; vulnerary; anticonvulsant; antibacterial;
 CC antifungal; antiparasitic and cardiant activity. The polynucleotide and
 CC protein sequences are used in the diagnosis of cancer, particularly
 CC breast and ovarian cancer. The nucleic acid sequences, proteins, agonists
 CC and antagonists may also be used in the diagnosis, prevention and treatment
 CC of immune disorders e.g. Addison's disease, allergies, autoimmune
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
 CC cardiovascular disorders such as myocardial ischaemias; wound healing;
 CC neurological diseases such as cerebral anoxia and epilepsy; and
 CC infectious diseases
 XX
 XX Sequence 102 AA;
 SQ
 Query Match 55.1%; Score 38; DB 3; Length 102;
 Best Local Similarity 53.8%; Pred. No. 1.6e+02;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 1 EGGPGTTSNRDLA 13
 DB 67 EGGPGAEGSRLLS 79
 RESULT 23
 AABG76902
 ID AABG76902 standard; protein; 134 AA.
 XX
 AC AABG76902;
 AC
 XX 05-NOV-2002 (first entry)
 DT
 XX Human fatty-acid binding protein-like protein.
 DE
 XX Human; NOVA; cardiomyopathy; atherosclerosis; cell signal processing;
 KW breast cancer; Alzheimer's disease; epilepsy; Huntington's disease;
 KW anxiety; behavioural disorder; multiple sclerosis; myasthenia gravis;
 KW neurodegeneration; Parkinson's disease; pain; stroke; endometriosis;
 KW autoimmune disease; allergy; addiction; asthma; transplantation;
 KW graft versus host disease; systemic lupus erythematosus; scleroderma;
 KW psoriasis; Crohn's disease; HIV infection; human immunodeficiency virus;

KW atherosclerosis; cirrhosis; rheumatoid arthritis; diabetes; pancreatitis;
 KW thrombocytopenia; bleeding disorder; metabolic disorder; obesity;
 KW glucose transport defect; glomerulonephritis; hypercalcaemia;
 KW polycystic kidney disease; renal tubular acidosis; skin disorder;
 KW congenital diarrhoea; respiratory disease; gastro-intestinal disease;
 KW muscle disorder; bone disorder; joint disorder; skeletal disorder;
 KW haematopoietic disorder; urinary system disorder; osteoporosis;
 KW dental disease; dental infection; growth disorder; reproductive disorder;
 KW hypogonadism; fertility disorder; viral infection; bacterial infection;
 KW parasitic infection; metabolic pathway modulation; gene therapy;
 KW zinc metallopeptase; ADAM-TS 7; alpha-2-macroglobulin precursor;
 KW ileal sodium/bile acid cotransporter; proinhibin; WT; CIP4; spinosin;
 KW macrophage stimulating protein precursor; fatty acid-binding protein;
 KW gap junction beta-5 protein; hepsin/plasma transmembrane serine protease.

XX Homo sapiens.

XX WO200233087-A2.

XX 25-APR-2002.

XX 17-OCT-2001; 2001WO-US032496.

XX 17-OCT-2000; 2000US-0241040P.

XX 17-OCT-2000; 2000US-0241058P.

XX 17-OCT-2000; 2000US-0241063P.

XX 20-OCT-2000; 2000US-0241243P.

XX 23-OCT-2000; 2000US-0242152P.

XX 23-OCT-2000; 2000US-0242482P.

XX 23-OCT-2000; 2000US-0242611P.

XX 23-OCT-2000; 2000US-0242612P.

XX 24-OCT-2000; 2000US-0242880P.

XX 24-OCT-2000; 2000US-0242881P.

XX 29-DEC-2000; 2000US-0259028P.

XX 20-FEB-2001; 2001US-0269813P.

XX 25-APR-2001; 2001US-0286324P.

XX 29-MAY-2001; 2001US-0294108P.

XX 09-JUL-2001; 2001US-0303698P.

XX 16-OCT-2001; 2001US-00981151.

XX (CURA-) CURAGEN CORP.

XX Edinger S, Gerlach V, Macdougall JR, Malyankar UM, Smithson G;

PI Millet I, Peyman JA, Stone DJ, Gunther E, Ellerman K, Shimkets RA;

PI Fadigar M, Guo X, Patturajan M, Taupier RJ, Burgess CE;

PI Zernhusen BD, Kekuda R, Spytek KA, Gangolli EA, Fernandes ER;

PI Gorman L;

XX WPI: 2002-590434/63.

DR N-PSDB; ABS59331.

XX Cytoplasmic, nuclear, membrane bound and secreted polypeptides and
 PT nucleic acids encoding the polypeptides for diagnosing and treating e.g.
 PT cancer, Alzheimer's disease, cardiomyopathy, metabolic disease and
 PT diabetes.

XX Claim 1; Page 78; 305pp; English.

XX The present invention relates to new NOVX (NOV1-10) polypeptides. The
 CC molecules of the invention are useful for treating or preventing a NOVX-
 CC associated disorder, such as cardiomyopathy, atherosclerosis, or a
 CC disorder related to cell signal processing and metabolic pathway
 CC modulation in humans. NOVX polypeptides, nucleic acids and antibodies are
 CC useful for treating or preventing disorders or syndromes including breast
 CC cancer, Alzheimer's disease, epilepsy, Huntington's disease, anxiety,
 CC behavioural disorders, multiple sclerosis, myasthenia gravis,
 CC neurodegeneration, Parkinson's disease, pain, stroke, autoimmune disease,
 CC allergies, addiction, asthma, endometriosis, graft versus host disease,
 CC systemic lupus erythematosus, scleroderma, transplantation, psoriasis,
 CC Crohn's disease, HIV (human immunodeficiency virus) infection,
 CC atherosclerosis, cirrhosis, rheumatoid arthritis, diabetes,
 CC thrombocytopenia, bleeding disorders, metabolic disorders, obesity,
 CC glucose transport defect, glomerulonephritis, hypercalcaemia, polycystic

CC kidney disease, pancreatitis, renal tubular acidosis, skin disorders,
 CC congenital diarrhoea, respiratory disease, gastro-intestinal diseases,
 CC muscle, bone, joint and skeletal disorders, haematopoietic disorders,
 CC urinary system disorders, osteoporosis, dental disease and infection,
 CC growth and reproductive disorders, hypogonadism, fertility, and/or other
 CC pathologies and disorders, viral, bacterial, or parasitic infections. The
 CC present amino acid sequence represents a NOVX protein of the invention

XX Sequence 134 AA;

Query Match 55.1%; Score 38; DB 5; Length 134;

Best Local Similarity 53.8%; Pred. No. 2.2e+02;

Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 EGGPGTTSNRLDA 13

Db 25 EGGVGTALRKMDA 37

RESULT 24

AAG24538

ID AAG24538 standard; protein; 174 AA.

XX AC AAG24538;

XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 28250.

XX Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-00301439.

XX 25-FEB-1999; 99US-0121825P.

XX 05-MAR-1999; 99US-0123180P.

XX 09-MAR-1999; 99US-0123548P.

XX 23-MAR-1999; 99US-0125788P.

XX 29-MAR-1999; 99US-0126264P.

XX 01-APR-1999; 99US-0126785P.

XX 06-APR-1999; 99US-0127462P.

XX 08-APR-1999; 99US-0128234P.

XX 16-APR-1999; 99US-0128714P.

XX 19-APR-1999; 99US-0129845P.

XX 21-APR-1999; 99US-0130077P.

XX 23-APR-1999; 99US-0130449P.

XX 28-APR-1999; 99US-0130510P.

XX 30-APR-1999; 99US-0130891P.

XX 04-MAY-1999; 99US-0131448P.

XX 06-MAY-1999; 99US-0132407P.

XX 06-MAY-1999; 99US-0132485P.

XX 07-MAY-1999; 99US-0132486P.

XX 11-MAY-1999; 99US-0132863P.

XX 14-MAY-1999; 99US-0134218P.

XX 14-MAY-1999; 99US-0134219P.

XX 14-MAY-1999; 99US-0134221P.

XX 18-MAY-1999; 99US-0134370P.

XX 19-MAY-1999; 99US-0134768P.

XX 20-MAY-1999; 99US-0134941P.

XX 21-MAY-1999; 99US-0135124P.

XX 24-MAY-1999; 99US-0135353P.

XX 24-MAY-1999; 99US-0135629P.

PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 21-JUN-1999; 99US-0139817P.
PR 21-JUN-1999; 99US-0139763P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144326P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 18-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.

```
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.
Query Match 55.1%; Score 38; DB 3; Length 174;
Best Local Similarity 50.0%; Pred. NO. 2.9e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 2 GPGGTTSNRLDA 13
DB ||||:|:|:
25 GPGGLSGNLIIDS 36
RESULT 25
AAG24537
ID. AAG24537 standard; protein; 184 AA.
XX
AC AAG24537;
XX
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 28249.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX PD 06-SEP-2000.
XX
XX PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134321P.
PR 18-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 21-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
```

PR 09-AUG-1999; 99US-0147493P.
 PR 09-AUG-1999; 99US-0147935P.
 PR 10-AUG-1999; 99US-0148171P.
 PR 11-AUG-1999; 99US-0148319P.
 PR 12-AUG-1999; 99US-0148341P.
 PR 13-AUG-1999; 99US-0148565P.
 PR 13-AUG-1999; 99US-0148684P.
 PR 16-AUG-1999; 99US-0149336P.
 PR 17-AUG-1999; 99US-0149175P.
 PR 18-AUG-1999; 99US-0149426P.
 PR 20-AUG-1999; 99US-0149722P.
 PR 20-AUG-1999; 99US-0149723P.
 PR 20-AUG-1999; 99US-0149929P.
 PR 23-AUG-1999; 99US-0149902P.
 PR 23-AUG-1999; 99US-0149930P.
 PR 25-AUG-1999; 99US-0150566P.
 PR 26-AUG-1999; 99US-0150884P.
 PR 27-AUG-1999; 99US-0151065P.
 PR 27-AUG-1999; 99US-0151066P.
 PR 27-AUG-1999; 99US-0151080P.
 PR 30-AUG-1999; 99US-0151303P.
 PR 31-AUG-1999; 99US-0151438P.
 PR 01-SEP-1999; 99US-0151930P.
 PR 07-SEP-1999; 99US-0152363P.
 PR 10-SEP-1999; 99US-0153070P.
 PR 13-SEP-1999; 99US-0153758P.
 PR 15-SEP-1999; 99US-0154018P.
 PR 16-SEP-1999; 99US-0154039P.
 PR 20-SEP-1999; 99US-0154779P.
 PR 22-SEP-1999; 99US-0155139P.
 PR 23-SEP-1999; 99US-0155486P.
 PR 24-SEP-1999; 99US-0155659P.
 PR 28-SEP-1999; 99US-0156458P.
 PR 29-SEP-1999; 99US-0156596P.
 PR 04-OCT-1999; 99US-0157117P.
 PR 05-OCT-1999; 99US-0157753P.
 PR 06-OCT-1999; 99US-0157865P.
 PR 07-OCT-1999; 99US-0158029P.
 PR 08-OCT-1999; 99US-0158232P.
 PR 12-OCT-1999; 99US-0158369P.
 PR 13-OCT-1999; 99US-0159293P.
 PR 13-OCT-1999; 99US-0159294P.
 PR 13-OCT-1999; 99US-0159295P.
 PR 14-OCT-1999; 99US-0159329P.
 PR 14-OCT-1999; 99US-0159330P.
 PR 14-OCT-1999; 99US-0159331P.
 PR 14-OCT-1999; 99US-0159637P.
 PR 18-OCT-1999; 99US-0159638P.
 PR 18-OCT-1999; 99US-0159584P.
 PR 21-OCT-1999; 99US-0160741P.
 PR 21-OCT-1999; 99US-0160767P.
 PR 21-OCT-1999; 99US-0160768P.
 PR 21-OCT-1999; 99US-0160770P.
 PR 21-OCT-1999; 99US-0160814P.
 PR 21-OCT-1999; 99US-0160815P.
 PR 22-OCT-1999; 99US-0160980P.
 PR 22-OCT-1999; 99US-0160981P.
 PR 22-OCT-1999; 99US-0160989P.
 PR 25-OCT-1999; 99US-0161404P.
 PR 25-OCT-1999; 99US-0161405P.
 PR 25-OCT-1999; 99US-0161406P.
 PR 26-OCT-1999; 99US-0161359P.
 PR 26-OCT-1999; 99US-0161360P.
 PR 26-OCT-1999; 99US-0161361P.
 PR 28-OCT-1999; 99US-0161920P.
 PR 28-OCT-1999; 99US-0161992P.
 PR 28-OCT-1999; 99US-0161993P.
 PR 29-OCT-1999; 99US-0162142P.

Query Match 55.1%; Score 38; DB 3; Length 184;
 Best Local Similarity 50.0%; Pred. NO. 3e+02;
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps. 0;

QY 2 GPGTTSNRDLA 13
 Db 35 GPGLSGNLIDS 46

Search completed: September 24, 2004, 01:53:31
 Job time : 94.6415 secs

This Page Blank (uspto)

Result No.	Score	Query		Length	DB	ID	Description
		Match	Time				
1	41	59.4	305	4	US-09-252-991A-22845		Sequence 22845, A
2	39	56.5	351	4	US-09-252-991A-17299		Sequence 17299, A
3	39	56.5	461	4	US-09-134-000C-5551		Sequence 5551, Ap
4	38	55.1	440	4	US-09-252-991A-27594		Sequence 27594, A
5	38	55.1	610	4	US-09-252-991A-20299		Sequence 20299, A
6	38	55.1	818	4	US-09-252-991A-16691		Sequence 16691, A
7	37	53.6	198	3	US-09-188-930-138		Sequence 138, App
8	37	53.6	188	4	US-09-312-283C-138		Sequence 138, App
9	37	53.6	285	4	US-09-312-283C-382		Sequence 382, App
10	37	53.6	285	4	US-09-552-204A-2		Sequence 2, Appli
11	37	53.6	294	3	US-09-188-930-294		Sequence 294, App
12	37	53.6	294	4	US-09-312-283C-294		Sequence 294, App
13	37	53.6	655	1	US-07-736-178C-2		Sequence 2, Appli
14	37	53.6	1003	4	US-09-252-991A-24069		Sequence 24069, A
15	37	53.6	1287	4	US-09-252-991A-29606		Sequence 29606, A
16	36	52.2	159	3	US-08-415-655-7		Sequence 7, Appli
17	36	52.2	229	4	US-09-134-000C-4443		Sequence 4443, Ap
18	36	52.2	266	4	US-09-199-637A-157		Sequence 157, App
19	36	52.2	275	4	US-09-252-991A-23227		Sequence 23227, A
20	36	52.2	300	3	US-09-064-703-10		Sequence 10, Appl
21	36	52.2	305	3	US-09-264-703-11		Sequence 11, Appl
22	36	52.2	325	4	US-09-252-991A-23098		Sequence 23098, A
23	36	52.2	417	3	US-08-640-906-4		Sequence 4, Appli
24	36	52.2	417	3	US-08-640-906-18		Sequence 18, Appl
25	36	52.2	417	4	US-09-395-936-4		Sequence 4, Appli
26	36	52.2	417	4	US-09-395-936-18		Sequence 18, Appl
27	36	52.2	420	4	US-09-252-991A-26653		Sequence 26653, A

```

101 34 49.3 735 4 US-09-252-991A-32240
102 34 49.3 1321 2 US-08-317-310A-64
103 34 49.3 1841 2 US-08-804-227C-6
104 34 49.3 1864 2 US-08-804-227C-3
105 34 49.3 4630 3 US-09-091-609-2
106 34 49.3 5215 3 US-09-105-537-2
107 33 47.8 50 4 US-09-621-976-6633
108 33 47.8 139 3 US-08-685-808-5
109 33 47.8 139 3 US-08-505-860C-5
110 33 47.8 139 3 US-09-252-991A-18730
111 33 47.8 176 4 US-09-252-991A-30542
112 33 47.8 178 4 US-09-252-991A-28710
113 33 47.8 191 1 US-08-709-912-10
114 33 47.8 191 2 US-09-047-370-10
115 33 47.8 201 4 US-09-252-991A-19786
116 33 47.8 202 4 US-09-252-991A-28470
117 33 47.8 206 1 US-08-315-695-19
118 33 47.8 215 1 US-08-044-621D-34
119 33 47.8 225 1 US-08-290-979A-8
120 33 47.8 225 4 US-09-570-856B-26
121 33 47.8 226 4 US-09-367-891A-2
122 33 47.8 235 4 US-09-328-352-5715
123 33 47.8 244 4 US-09-252-991A-18465
124 33 47.8 256 4 US-09-134-001C-3271
125 33 47.8 261 4 US-09-252-991A-23810
126 33 47.8 264 1 US-08-107-042-2
127 33 47.8 268 2 US-08-824-874-1
128 33 47.8 268 3 US-09-210-084-1
129 33 47.8 268 4 US-09-764-762-1
130 33 47.8 287 4 US-09-252-991A-26715
131 33 47.8 313 4 US-09-252-991A-20483
132 33 47.8 325 4 US-09-252-991A-28580
133 33 47.8 335 4 US-09-570-856B-15
134 33 47.8 336 4 US-09-394-455-2
135 33 47.8 337 4 US-09-198-452A-411
136 33 47.8 343 4 US-09-394-455-15
137 33 47.8 343 4 US-09-394-455-34
138 33 47.8 343 4 US-09-394-455-34
139 33 47.8 344 4 US-09-540-236-3062
140 33 47.8 349 4 US-09-543-681A-6158
141 33 47.8 350 4 US-09-457-040B-37
142 33 47.8 351 4 US-09-457-040B-6
143 33 47.8 351 4 US-09-394-455-4
144 33 47.8 361 4 US-09-339-159B-30
145 33 47.8 387 4 US-09-252-991A-27685
146 33 47.8 405 1 US-08-406-070-2
147 33 47.8 406 3 US-09-070-356-2
148 33 47.8 413 2 US-08-481-814A-8
149 33 47.8 413 3 US-08-836-582-2
150 33 47.8 413 4 US-09-265-566-2

```

ALIGNMENTS

```

RESULT 1
US-09-252-991A-22845
; Sequence 22845, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22845
; LENGTH: 305

```

```

; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22845
Query Match 59.4%; Score 41; DB 4; Length 305;
Best Local Similarity 58.3%; Pred. No. 34;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EGGPGTTSNRLD 12
   : ||||| : |||
Db 224 QGAPGTGASRLD 235

RESULT 2
US-09-252-991A-17299
; Sequence 17299, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17299
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17299
Query Match 56.5%; Score 39; DB 4; Length 351;
Best Local Similarity 77.8%; Pred. No. 86;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GGGPGTTSNR 10
   ||||| ||
Db 14 GGGPGTLNR 22

RESULT 3
US-09-134-000C-5551
; Sequence 5551, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5551
; LENGTH: 461
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-5551
Query Match 56.5%; Score 39; DB 4; Length 461;
Best Local Similarity 80.0%; Pred. No. 1-2e+02;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GGGPGTTSNR 11
   ||||| |||
Db 159 GRPGTFSNRL 168

```

```
RESULT 4
US-09-252-991A-27594
; Sequence 27594, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27594
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27594

Query Match 55.1%; Score 38; DB 4; Length 440;
Best Local Similarity 70.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 GGGPGTTSNRL 11
DB 263 GGGPGPARNRL 272

RESULT 5
US-09-252-991A-20299
; Sequence 20299, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20299
; LENGTH: 610
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20299

Query Match 55.1%; Score 38; DB 4; Length 610;
Best Local Similarity 58.3%; Pred. No. 2.3e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EGGPGTTSNRLD 12
DB 394 EGGPDTPHRL 405

RESULT 6
US-09-252-991A-16691
; Sequence 16691, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
```

```
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 16691
; LENGTH: 818
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-16691

Query Match 55.1%; Score 38; DB 4; Length 818;
Best Local Similarity 61.5%; Pred. No. 3.2e+02;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 EGGPGTTSNRLDA 13
DB 718 EGGAGEQNARLDA 730

RESULT 7
US-09-188-930-138
; Sequence 138, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.1011c1
; CURRENT APPLICATION NUMBER: US/09/188,930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 138
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Rat
US-09-188-930-138

Query Match 53.6%; Score 37; DB 3; Length 198;
Best Local Similarity 70.0%; Pred. No. 98;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EGGPGTTSNR 10
DB 90 EGGPGRTGNR 99

RESULT 8
US-09-312-283C-138
; Sequence 138, Application US/09312283C
; Patent No. 6573095
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated from Skin Cells
; FILE REFERENCE: 11000.1011c2
; CURRENT APPLICATION NUMBER: US/09/312,283C
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 425
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 138
; LENGTH: 198
```

```
; TYPE: PRT
; ORGANISM: Mouse
US-09-312-283C-138

Query Match          53.6%; Score 37; DB 4; Length 198;
Best Local Similarity 70.0%; Pred. No. 98;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EGGPGTTSNR 10
Db 90 EGGPGRTGNR 99

RESULT 9
US-09-312-283C-382
; Sequence 382, Application US/09312283C
; Patent No. 6573095
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated from Skin Cells
; TITLE OF INVENTION: and Methods For Their Use
; FILE REFERENCE: 11000.1011c2
; CURRENT APPLICATION NUMBER: US/09/312,283C
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 425
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 382
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Mouse
US-09-312-283C-382

Query Match          53.6%; Score 37; DB 4; Length 285;
Best Local Similarity 70.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EGGPGTTSNR 10
Db 81 EGGPGRTGNR 90

RESULT 10
US-09-552-204A-2
; Sequence 2, Application US/09552204A
; Patent No. 6620909
; GENERAL INFORMATION:
; APPLICANT: Piddington, Christopher S.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOG ZACRP2
; FILE REFERENCE: 99-08
; CURRENT APPLICATION NUMBER: US/09/552,204A
; CURRENT FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: 60/130,207
; PRIOR FILING DATE: 1999-04-20
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-552-204A-2

Query Match          53.6%; Score 37; DB 4; Length 285;
Best Local Similarity 70.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EGGPGTTSNR 10
Db 90 EGGPGRTGNR 99

RESULT 11
US-09-188-930-294
; Sequence 294, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods For Their Use
; FILE REFERENCE: 11000.1011c1
; CURRENT APPLICATION NUMBER: US/09/188,930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 294
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Rat
US-09-188-930-294

Query Match          53.6%; Score 37; DB 3; Length 294;
Best Local Similarity 70.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EGGPGTTSNR 10
Db 90 EGGPGRTGNR 99

RESULT 12
US-09-312-283C-294
; Sequence 294, Application US/09312283C
; Patent No. 6573095
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated from Skin Cells
; TITLE OF INVENTION: and Methods for Their Use
; FILE REFERENCE: 11000.1011c2
; CURRENT APPLICATION NUMBER: US/09/312,283C
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 425
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 294
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Mouse
US-09-312-283C-294

Query Match          53.6%; Score 37; DB 4; Length 294;
Best Local Similarity 70.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EGGPGTTSNR 10
Db 90 EGGPGRTGNR 99

RESULT 13
US-07-736-178C-2
; Sequence 2, Application US/07736178C
; Patent No. 5514544
; GENERAL INFORMATION:
```

APPLICANT: Rao, Ramachandra N
APPLICANT: Turner, Jan R
TITLE OF INVENTION: ACTIVATOR GENE FOR MACROLIDE
TITLE OF INVENTION: BIOSYNTHESIS
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Eli Lilly And Company
STREET: Lilly Corporate Center
CITY: Indianapolis
STATE: IN
COUNTRY: USA
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION NUMBER: US/07/736,178C
FILING DATE: 19910726
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Conrad, Robert A
REGISTRATION NUMBER: 32089
REFERENCE/DOCKET NUMBER: X8144
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-3334
TELEFAX: 317-276-1294
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 655 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-736-178C-2

Query Match 53.6%; Score 37; DB 1; Length 655;
Best Local Similarity 72.7%; Pred. No. 3.7e+02;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 GGGGTTSNRLD 12
||| |||||
DB 436 GGGPGTSPRLD 446

RESULT 14
US-09-252-991A-24069
Sequence 24069, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 24069
LENGTH: 1003
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24069

Query Match 53.6%; Score 37; DB 4; Length 1003;
Best Local Similarity 75.0%; Pred. No. 5.9e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 PGTTSNRL 11
|||||:

Db 306 PGTANRM 313
RESULT 15
US-09-252-991A-29606
Sequence 29606, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 29606
LENGTH: 1287
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29606

Query Match 53.6%; Score 37; DB 4; Length 1287;
Best Local Similarity 63.6%; Pred. No. 7.7e+02;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 GGGGTTSNRLD 12
|||||:
DB 326 GGGGTTNGVMD 336

RESULT 16

US-08-415-655-7
Sequence 7, Application US/08415655
Patent No. 6025480
GENERAL INFORMATION:
APPLICANT: Massague, Joan
APPLICANT: Lee, Mong-hong
TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ENCODING
TITLE OF INVENTION: P7SKIP2, A CYCLIN-DEPENDENT KINASE INHIBITOR AND USES OF
TITLE OF INVENTION: SAME
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/415,655
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 1747/47418
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 159 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant

; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-415-655-7

Query Match 52.2%; Score 36; DB 3; Length 159;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GGGGTTSNR 10
|||||:
Db 127 GGGTTSQGR 135

RESULT 17
US-09-134-000C-4443
; Sequence 4443, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 4443
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-4443

Query Match 52.2%; Score 36; DB 4; Length 229;
Best Local Similarity 70.0%; Pred. No. 1.7e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 GPGTTSNR 12
|||||:
Db 53 GPGTTPGLD 62

RESULT 18
US-09-199-637A-157
; Sequence 157, Application US/09199637A
; Patent No. 6355411
; GENERAL INFORMATION:
; APPLICANT: Ausubel, Frederick
; APPLICANT: Goodman, Howard M.
; APPLICANT: Rahme, Laurence G.
; APPLICANT: Mahajan-Wikios, Shalina
; APPLICANT: Tan, Man-Wah
; APPLICANT: Cao, Hui
; APPLICANT: Drenkard, Eliana
; APPLICANT: Tsongalis, John
; TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID
; FILE OF INVENTION: SEQUENCES AND USES THEREOF
; FILE REFERENCE: 00786/361002
; CURRENT APPLICATION NUMBER: US/09/199,637A
; CURRENT FILING DATE: 1998-11-25
; PRIOR APPLICATION NUMBER: 60/066,517
; PRIOR FILING DATE: 1997-11-25
; NUMBER OF SEQ ID NOS: 437
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 157
; LENGTH: 266
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-199-637A-157

Query Match 52.2%; Score 36; DB 4; Length 266;
Best Local Similarity 54.5%; Pred. No. 2e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 GGGGTTSNR 12
|||||:
Db 158 GEPDTNNRME 168

RESULT 19
US-09-252-991A-29227
; Sequence 29227, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29227
; LENGTH: 275
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29227

Query Match 52.2%; Score 36; DB 4; Length 275;
Best Local Similarity 54.5%; Pred. No. 2.1e+02;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EGGPGTTSNR 11
:|||||:
Db 246 KGGPGTAENRI 256

RESULT 20
US-09-064-703-10
; Sequence 10, Application US/09064703
; Patent No. 6033894
; GENERAL INFORMATION:
; APPLICANT: Craik, Charles S.
; APPLICANT: Unal, Ayce
; APPLICANT: Ganem, Donald E.
; TITLE OF INVENTION: Kaposi's Syndrome Herpesvirus
; TITLE OF INVENTION: Protease and Assembly Protein Compositions and
; TITLE OF INVENTION: Methods
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: P.O. Box 60850
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/064,703
; FILING DATE: 22-APR-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/044,152
; FILING DATE: 22-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Petithory, Joanne R
; REGISTRATION NUMBER: P42,995


```

; APPLICANT: Blumenkopf, Todd A.
; APPLICANT: Cory, Michael
; TITLE OF INVENTION: Cell-Targeting Molecule Comprising a Mutant Human
; FILE REFERENCE: PB1500USW
; CURRENT APPLICATION NUMBER: US/08/640,906B
; CURRENT FILING DATE: 1996-05-09
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 417
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-640-906-18

Query Match      52.2%; Score 36; DB 3; Length 417;
Best Local Similarity 75.0%; Pred. No. 3.3e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      2 GPGTTSN 9
DB      260 GPGGASN 267

RESULT 25
US-09-395-936-4
; Sequence 4, Application US/09395936
; Patent No. 6319702
; GENERAL INFORMATION:
; APPLICANT: Smith, Gary
; APPLICANT: Cory, Michael
; APPLICANT: Blumenkopf, Todd
; TITLE OF INVENTION: Nucleic Acids Encoding Mutant Human Carboxypeptidase A
; FILE REFERENCE: PB1500US3
; CURRENT APPLICATION NUMBER: US/09/395,936
; CURRENT FILING DATE: 1999-09-14
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 417
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-395-936-4

Query Match      52.2%; Score 36; DB 4; Length 417;
Best Local Similarity 75.0%; Pred. No. 3.3e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      2 GPGTTSN 9
DB      260 GPGGASN 267

Search completed: September 24, 2004, 02:20:06
Job time : 25.0755 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 01:54:43 ; Search time 75.3019 Seconds
(without alignments)
55.513 Million cell updates/sec

Title: US-10-088-639A-2_COPY_226_238

Perfect score: 69

Sequence: 1 EGGPGTTSNRLDA 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1349238 seqs, 321558718 residues

Total number of hits satisfying chosen parameters: 1349238

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/prodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/prodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/prodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/prodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/prodata/2/pubpaa/US07_PUBCOMB.pep.*
- 6: /cgn2_6/prodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/prodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/prodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/prodata/2/pubpaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/prodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/prodata/2/pubpaa/US09_PUBCOMB.pep.*
- 12: /cgn2_6/prodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/prodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/prodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/prodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/prodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/prodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/prodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	63.8	3149	16	US-10-408-765A-2326
2	42	60.9	566	12	US-10-424-599-166000
3	41	59.4	398	15	US-10-369-493-4888
4	41	59.4	400	15	US-10-369-493-7646
5	41	59.4	19608	15	US-10-084-846A-8
6	39	56.5	69	12	US-10-424-599-163360
7	39	56.5	85	12	US-10-424-599-206460
8	39	56.5	130	12	US-10-282-122A-69481
9	39	56.5	208	16	US-10-767-701-39568
10	39	56.5	203	16	US-10-437-963-190792
11	39	56.5	219	9	US-09-893-737-130
12	39	56.5	385	14	US-10-156-761-12542
13	39	56.5	454	9	US-09-815-242-10700
14	39	56.5	454	12	US-10-282-122A-57133
15	39	56.5	600	12	US-10-182-243-45
					Sequence 2326, Ap
					Sequence 166000,
					Sequence 4888, Ap
					Sequence 7646, Ap
					Sequence 8, Appli
					Sequence 163360,
					Sequence 206460,
					Sequence 69481, A
					Sequence 39568, A
					Sequence 190792,
					Sequence 130, App
					Sequence 12542, A
					Sequence 10700, A
					Sequence 57133, A
					Sequence 45, Appl

Sequence 18, Appl
Sequence 797, App
Sequence 797, App
Sequence 141326,
Sequence 786, App
Sequence 786, App
Sequence 18, Appl
Sequence 24992,
Sequence 195589,
Sequence 38825, A
Sequence 65647, A
Sequence 1149, Ap
Sequence 159302,
Sequence 14175, A
Sequence 3298, Ap
Sequence 8193, Ap
Sequence 47371, A
Sequence 43208, A
Sequence 150859,
Sequence 122878,
Sequence 171832,
Sequence 2611, Ap
Sequence 67846, A
Sequence 138, App
Sequence 382, App
Sequence 88, Appl
Sequence 5, Appli
Sequence 2, Appli
Sequence 294, App
Sequence 8, Appli
Sequence 3592, Ap
Sequence 2619, Ap
Sequence 6, Appli
Sequence 30, Appl
Sequence 354, App
Sequence 951, App
Sequence 51, Appl
Sequence 9, Appli
Sequence 3903, Ap
Sequence 4, Appli
Sequence 26, Appl
Sequence 4, Appli
Sequence 11, Appl
Sequence 57604, A
Sequence 112022,
Sequence 219730,
Sequence 189212,
Sequence 201467,
Sequence 236779,
Sequence 53, Appl
Sequence 55, Appl
Sequence 70, Appl
Sequence 71, Appl
Sequence 72, Appl
Sequence 9, Appli
Sequence 10, Appl
Sequence 13, Appl
Sequence 14, Appl
Sequence 17, Appl
Sequence 18, Appl
Sequence 21, Appl
Sequence 22, Appl
Sequence 53897, A
Sequence 102592,
Sequence 1198, Ap
Sequence 1012, Ap
Sequence 69656, A
Sequence 162296,
Sequence 60885, A
Sequence 183411,

89 36 52.2 222 12 US-10-424-599-208766
90 36 52.2 226 12 US-10-424-599-238520
91 36 52.2 256 14 US-10-301-822-65
92 36 52.2 256 15 US-10-094-749-2770
93 36 52.2 265 15 US-10-094-749-2763
94 36 52.2 266 10 US-09-975-719-157
95 36 52.2 286 12 US-10-307-817-246
96 36 52.2 328 16 US-10-437-963-181524
97 36 52.2 406 12 US-10-307-817-238
98 36 52.2 410 15 US-10-369-493-366
99 36 52.2 414 12 US-10-307-817-242
100 36 52.2 414 12 US-10-307-817-244
101 36 52.2 417 9 US-09-923-779-150
102 36 52.2 417 12 US-10-307-817-228
103 36 52.2 417 12 US-10-307-817-250
104 36 52.2 417 12 US-10-307-817-576
105 36 52.2 417 12 US-10-307-817-577
106 36 52.2 417 12 US-10-307-817-579
107 36 52.2 418 12 US-10-307-817-234
108 36 52.2 418 12 US-10-307-817-248
109 36 52.2 418 12 US-10-307-817-572
110 36 52.2 418 12 US-10-307-817-578
111 36 52.2 419 12 US-10-307-817-230
112 36 52.2 424 12 US-10-307-817-236
113 36 52.2 425 12 US-10-307-817-226
114 36 52.2 425 12 US-10-307-817-232
115 36 52.2 425 12 US-10-307-817-574
116 36 52.2 425 12 US-10-307-817-580
117 36 52.2 428 9 US-09-925-297-528
118 36 52.2 428 12 US-10-424-599-238521
119 36 52.2 473 12 US-10-424-599-153291
120 36 52.2 473 14 US-10-029-386-33370
121 36 52.2 538 12 US-10-282-122A-60329
122 36 52.2 545 16 US-10-437-963-154817
123 36 52.2 553 12 US-10-424-599-226659
124 36 52.2 558 12 US-10-282-122A-48557
125 36 52.2 575 14 US-10-372-095-5
126 36 52.2 603 16 US-10-437-963-169783
127 36 52.2 620 10 US-09-931-009A-1
128 36 52.2 644 10 US-09-919-039-121
129 36 52.2 644 12 US-10-236-392-32
130 36 52.2 661 12 US-10-425-114-64386
131 36 52.2 696 16 US-10-437-963-132807
132 36 52.2 831 15 US-10-360-101-236
133 36 52.2 847 13 US-10-112-527-4
134 36 52.2 884 12 US-10-282-122A-49820
135 36 52.2 1132 16 US-10-437-963-120040
136 36 52.2 1203 12 US-09-769-699-2
137 36 52.2 1294 16 US-10-437-963-192226
138 36 52.2 1592 16 US-10-437-963-120039
139 36 52.2 2436 9 US-09-795-693-8
140 36 52.2 2436 14 US-10-156-239-8
141 36 52.2 2436 14 US-10-199-485-8
142 36 52.2 2436 16 US-10-380-727-2
143 35.5 51.4 793 14 US-10-156-761-8812
144 35 50.7 25 10 US-09-764-891-3390
145 35 50.7 56 9 US-09-864-761-44013
146 35 50.7 77 16 US-10-437-963-199907
147 35 50.7 78 10 US-09-731-872-390
148 35 50.7 78 10 US-09-876-997-390
149 35 50.7 80 9 US-09-867-550-1672
150 35 50.7 80 11 US-09-833-245-1916

ALIGNMENTS

RESULT 1

US-10-408-765A-2326
; Sequence 2326, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Wartock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2326
; LENGTH: 3149
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-2326

Query Match 63.8%; Score 44; DB 16; Length 3149;
Best Local Similarity 80.0%; Pred. No. 8.5e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 GPGTTSNRLLD 12
|||:|||||
DB 72 GPGTSRRLLD 81

RESULT 2

US-10-424-599-166000
; Sequence 166000, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 166000
; LENGTH: 566
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_120914C.1.pap
US-10-424-599-166000

Query Match 60.9%; Score 42; DB 12; Length 566;
Best Local Similarity 72.7%; Pred. No. 3e+02;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 GPGTTSNRLLD 12
|||:|||||
DB 549 GPGTTSRRLLD 559

RESULT 3

US-10-369-493-4888
; Sequence 4888, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B

; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 4888
; LENGTH: 398
; TYPE: PRT
; ORGANISM: Burkholderia fungorum
US-10-369-493-4888

Query Match 59.4%; Score 41; DB 15; Length 398;
Best Local Similarity 72.7%; Pred. No. 3e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EGGPGTTSNRL 11
: | | | | |
DB 77 QGVPGTTVNR 87

RESULT 4
US-10-369-493-7646
; Sequence 7646, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 7646
; LENGTH: 400
; TYPE: PRT
; ORGANISM: Burkholderia cepacia
US-10-369-493-7646

Query Match 59.4%; Score 41; DB 15; Length 400;
Best Local Similarity 72.7%; Pred. No. 3e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EGGPGTTSNRL 11
: | | | | |
DB 79 QGVPGTTVNR 89

RESULT 5
US-10-084-846A-8
; Sequence 8, Application US/10084846A
; Publication No. US2004006026A1
; GENERAL INFORMATION:
; APPLICANT: WEITNAUER, GABRIELE
; APPLICANT: MUHLENWEG, AGNES
; APPLICANT: TREPFER, AXEL
; APPLICANT: BECHTHOLD, ANDREAS
; TITLE OF INVENTION: AVILAMYCIN DERIVATIVES
; FILE REFERENCE: 1974-005
; CURRENT APPLICATION NUMBER: US/10/084,846A
; CURRENT FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: PCT/EP01/09815
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: DE 101 09 166.4
; PRIOR FILING DATE: 2001-02-25
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 8

; LENGTH: 19608
; TYPE: PRT
; ORGANISM: Streptomyces viridochromogenes
; FEATURE:
; OTHER INFORMATION: Protein 3: amino acid sequence encoded by coding strand 1.
; OTHER INFORMATION: Start codon: atc, Start position: nucleotide 3.
US-10-084-846A-8

Query Match 59.4%; Score 41; DB 15; Length 19608;
Best Local Similarity 66.7%; Pred. No. 1.7e+04;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 GGPQTTSNRLDA 13
: | | | | |
DB 4261 GGPQTTSNRLDA 4272

RESULT 6
US-10-424-599-163360
; Sequence 163360, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 163360
; LENGTH: 69
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_118532C.1.pep
US-10-424-599-163360

Query Match 56.5%; Score 39; DB 12; Length 69;
Best Local Similarity 63.6%; Pred. No. 99;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EGGPGTTSNRL 11
: | | | | |
DB 42 EGGPQTSNRM 52

RESULT 7
US-10-424-599-206460
; Sequence 206460, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 206460
; LENGTH: 85
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_28460C.1.pep
US-10-424-599-206460

Query Match 56.5%; Score 39; DB 12; Length 85;

Best Local Similarity 87.5%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GGGGTTSN 9
|||
Db 52 GGGGTTIN 59

RESULT 8

US-10-282-122A-69481
; Sequence 69481, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 69481
; LENGTH: 130
; TYPE: PRT
; ORGANISM: Pseudomonas syringae
US-10-282-122A-69481

Query Match 56.5%; Score 39; DB 12; Length 130;
Best Local Similarity 61.5%; Pred. No. 1.9e+02;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 EGGGTTSNRLDA 13
|||
Db 11 EGGPGSTPRRLGA 23

RESULT 9

US-10-767-701-39568
; Sequence 39568, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(533535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 39568
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C20858_1.pep
US-10-767-701-39568

Query Match 56.5%; Score 39; DB 16; Length 188;
Best Local Similarity 63.6%; Pred. No. 2.8e+02;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 GPGTTSNRLDA 13
|||
Db 109 GPGTTDEKIDA 119

RESULT 10

US-10-437-963-190792
; Sequence 190792, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 190792
; LENGTH: 203
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_87172C.1.psp
US-10-437-963-190792

Query Match 56.5%; Score 39; DB 16; Length 203;
Best Local Similarity 70.0%; Pred. No. 3e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 GPGTTSNRLD 12
|||
Db 71 GPGTSSHRSD 80

RESULT 11

US-09-893-737-130
; Sequence 130, Application US/09893737
; Patent No. US20020110855A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Presnell, Scott R.
; TITLE OF INVENTION: MAMMALIAN SECRETED PROTEINS
; FILE REFERENCE: 00-41
; CURRENT APPLICATION NUMBER: US/09/893,737
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US 60/215,446

;
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 329
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 130
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-893-737-130

Query Match 56.5%; Score 39; DB 9; Length 219;
Best Local Similarity 46.2%; Pred. No. 3.3e+02;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 EGGPGTTSNRLDA 13
Db 33 QGGPGSTNSKROA 45
:||||:|

RESULT 12

US-10-156-761-12542
; Sequence 12542, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 12542
; LENGTH: 385
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-12542

Query Match 56.5%; Score 39; DB 14; Length 385;
Best Local Similarity 77.8%; Pred. No. 5.9e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 GGGPGTTSNR 10
Db 364 GGGPGTSSR 372
|||||

RESULT 13

US-09-815-242-10700
; Sequence 10700, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: Prokaryotes
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078

;
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10700
; LENGTH: 454
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-815-242-10700

Query Match 56.5%; Score 39; DB 9; Length 454;
Best Local Similarity 80.0%; Pred. No. 7e+02;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GGGGTTSNRL 11
Db 152 GRPGTFSNRL 161
|||||

RESULT 14

US-10-282-122A-57133
; Sequence 57133, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 57133
; LENGTH: 454
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-10-282-122A-57133

Query Match 56.5%; Score 39; DB 12; Length 454;
Best Local Similarity 80.0%; Pred. No. 7e+02;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GGGTTSNRL 11
Db 152 GRGTFNRL 161

RESULT 15
US-10-182-243-45
; Sequence 45, Application US/10182243
; Publication No. US20040048310A1
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY D.
; APPLICANT: WHYTE, DAVID
; APPLICANT: MANNING, GERARD
; APPLICANT: SUDARSANAM, SUCHA
; APPLICANT: MARTINEZ, RICARDO
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES AND PROTEIN KINASE-LIKE
; TITLE OF INVENTION: ENZYMES
; FILE REFERENCE: 038602/1366
; CURRENT APPLICATION NUMBER: US/10/182,243
; CURRENT FILING DATE: 2003-07-07
; PRIOR APPLICATION NUMBER: PCT/US01/02337
; PRIOR FILING DATE: 2001-01-25
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 45
; LENGTH: 600
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-182-243-45

Query Match 56.5%; Score 39; DB 12; Length 600;
Best Local Similarity 87.5%; Pred. No. 9.4e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EGGPGTTS 8
Db 187 EGGPGGATS 194

RESULT 16
US-10-415-011-18
; Sequence 18, Application US/10415011
; Publication No. US20040053394A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION
; APPLICANT: GURURAJAN, Rajagopal
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: CHAWLA, Narinder K.
; APPLICANT: ELLIOTT, Vicki S.
; APPLICANT: XU, Yuming
; APPLICANT: ARVIZU, Chandra S.
; APPLICANT: YAO, Monique G.
; APPLICANT: RAMKUMAR, Jayalaxmi
; APPLICANT: DING, Li
; APPLICANT: TANG, Y. Tom
; APPLICANT: HAFALIA, April J.A.
; APPLICANT: NGUYEN, Dannel B.
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: LU, Yan
; APPLICANT: YUE, Henry
; APPLICANT: BURFORD, Neil
; APPLICANT: BANDMAN, Olga

; APPLICANT: TRIBOULEY, Catherine M.
; APPLICANT: LAL, Preeti G.
; APPLICANT: RECIPON, Shirley A.
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: BOROWSKY, Mark L.
; APPLICANT: THORNTON, Michael B.
; APPLICANT: SWARNAKER, Anita
; APPLICANT: THANGAVELU, Kavitha
; APPLICANT: KHAN, Farrah A.
; APPLICANT: ISON, Craig H.
; TITLE OF INVENTION: HUMAN KINASES
; FILE REFERENCE: PI-0262 USN
; CURRENT APPLICATION NUMBER: US/10/415,011
; CURRENT FILING DATE: 2003-04-18
; PRIOR APPLICATION NUMBER: PCT/US01/47728
; PRIOR FILING DATE: 2001-10-20
; PRIOR APPLICATION NUMBER: US 60/242,410
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US 60/244,068
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: US 60/245,708
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: US 60/247,672
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: US 60/249,565
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: US 60/252,730
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: US 60/250,807
; PRIOR FILING DATE: 2000-12-01
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PERL Program
; SEQ ID NO 18
; LENGTH: 600
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20040053394A1 7474648CD1
US-10-415-011-18

Query Match 56.5%; Score 39; DB 12; Length 600;
Best Local Similarity 87.5%; Pred. No. 9.4e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EGGPGTTS 8
Db 187 EGGPGGATS 194

RESULT 17
US-09-925-302-797
; Sequence 797, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 797
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (170)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (446)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (506)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (577)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (583)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (584)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (599)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (584)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (608)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-09-925-302-797

Query Match 56.5%; Score 39; DB 9; Length 609;
Best Local Similarity 70.0%; Pred. No. 9.5e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GGPQTTSNRL 11
|||: |||
Db 586 GGPQSGGNRL 595

RESULT 18

US-09-925-302-797
; Sequence 797, Application US/09925302
; Publication No. US20030064072A9

; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 797
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:

; NAME/KEY: SITE
; LOCATION: (170)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (446)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (506)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (577)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (583)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (584)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; LOCATION: (599)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (608)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-09-925-302-797

Query Match 56.5%; Score 39; DB 12; Length 609;
Best Local Similarity 70.0%; Pred. No. 9.5e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GGPQTTSNRL 11
|||: |||
Db 586 GGPQSGGNRL 595

RESULT 19

US-10-437-963-141326
; Sequence 141326, Application US/10437963
; Publication No. US20040123343A1

; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 141326
; LENGTH: 97
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT4530_42440C.1.pap
US-10-437-963-141326

Query Match 55.1%; Score 38; DB 16; Length 97;
Best Local Similarity 53.8%; Pred. No. 2e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 EGGPGTTSNRLDA 13
|||: |||
Db 78 QGGPQSRQNRKDA 90

RESULT 20

US-09-925-298-786
; Sequence 786, Application US/09925298
; Publication No. US20020039764A1

; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103
; CURRENT APPLICATION NUMBER: US/09/925,298
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 786
; LENGTH: 102
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:

NAME/KEY: SITE
 LOCATION: (8)
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 NAME/KEY: SITE
 LOCATION: (33)
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 NAME/KEY: SITE
 LOCATION: (64)
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 NAME/KEY: SITE
 LOCATION: (83)
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 NAME/KEY: SITE
 LOCATION: (84)
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 NAME/KEY: SITE
 LOCATION: (86)
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 US-09-925-298-786

Query Match 55.1%; Score 38; DB 12; Length 102;
 Best Local Similarity 53.8%; Pred. No. 2.1e+02;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 EGGPGTTSNRLDA 13
 Db 67 EGGPGAEGSRLLDS 79

RESULT 21
 US-10-102-806-786
 ; Sequence 786, Application US/10102806
 ; Publication No. US20030054421A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
 ; FILE REFERENCE: P103P1C1
 ; CURRENT APPLICATION NUMBER: US/10/102,806
 ; PRIOR FILING DATE: 2002-03-22
 ; PRIOR APPLICATION NUMBER: 09/925,298
 ; PRIOR FILING DATE: 2001-08-10
 ; PRIOR APPLICATION NUMBER: PCT/US00/05881
 ; PRIOR FILING DATE: 2000-03-08
 ; PRIOR APPLICATION NUMBER: 60/124,270
 ; PRIOR FILING DATE: 1999-03-12
 ; NUMBER OF SEQ ID NOS: 846
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 786
 ; LENGTH: 102
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: SITE
 ; LOCATION: (8)
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 ; NAME/KEY: SITE
 ; LOCATION: (33)
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 ; NAME/KEY: SITE
 ; LOCATION: (64)
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 ; NAME/KEY: SITE
 ; LOCATION: (83)
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 ; NAME/KEY: SITE
 ; LOCATION: (84)
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 ; NAME/KEY: SITE
 ; LOCATION: (86)
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 ; US-10-102-806-786

Query Match 55.1%; Score 38; DB 14; Length 102;

Best Local Similarity 53.8%; Pred. No. 2.1e+02;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 1 EGGPGTTSNRLDA 13
 Db 67 EGGPGAEGSRLLDS 79

RESULT 22

US-09-981-151A-18
 ; Sequence 18, Application US/09981151A
 ; Publication No. US20030212256A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Edinger, Shlomit R
 ; APPLICANT: Gerlach, Valerie
 ; APPLICANT: MacDougall, John R
 ; APPLICANT: Malyankar, Muriel M
 ; APPLICANT: Smithson, Glenda
 ; APPLICANT: Millet, Isabelle
 ; APPLICANT: Feyman, John A
 ; APPLICANT: Stone, David J
 ; APPLICANT: Gunther, Erik
 ; APPLICANT: Ellerman, Karen
 ; APPLICANT: Shimkets, Richard A
 ; APPLICANT: Padigaru, Muralidhara
 ; APPLICANT: Guo, Xiaojia
 ; APPLICANT: Patturajan, Meera
 ; APPLICANT: Taupier Jr, Raymond J
 ; APPLICANT: Burgess, Catherine E
 ; APPLICANT: Zerhusen, Bryan D
 ; APPLICANT: Kekuda, Ramesh
 ; APPLICANT: Spytek, Kimberly A
 ; APPLICANT: Gangolli, Esha A
 ; APPLICANT: Fernandes, Elma R
 ; APPLICANT: Gorman, Linda
 ; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
 ; FILE REFERENCE: 21402-168
 ; CURRENT APPLICATION NUMBER: US/09/981,151A
 ; PRIOR FILING DATE: 2001-10-16
 ; PRIOR APPLICATION NUMBER: 60/241,040
 ; PRIOR FILING DATE: 2000-10-17
 ; PRIOR APPLICATION NUMBER: 60/241,058
 ; PRIOR FILING DATE: 2000-10-17
 ; PRIOR APPLICATION NUMBER: 60/241,063
 ; PRIOR FILING DATE: 2000-10-17
 ; PRIOR APPLICATION NUMBER: 60/241,243
 ; PRIOR FILING DATE: 2000-10-17
 ; PRIOR APPLICATION NUMBER: 60/242,152
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/242,482
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/242,611
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/242,612
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/242,880
 ; PRIOR FILING DATE: 2000-10-24
 ; PRIOR APPLICATION NUMBER: 60/242,881
 ; PRIOR FILING DATE: 2000-10-24
 ; Remaining prior application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 160
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 18
 ; LENGTH: 134
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-981-151A-18

Query Match 55.1%; Score 38; DB 11; Length 134;
 Best Local Similarity 53.8%; Pred. No. 2.8e+02;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 1 EGGPGTTSNRLDA 13

```
Db          25 ECGVGTALRKMDA 37
|||||
RESULT 23
US-10-424-599-248992
; Sequence 248992, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599.
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 248992
; LENGTH: 151
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(151)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_66870C.1.pep
US-10-424-599-248992

Query Match          55.1%; Score 38; DB 12; Length 151;
Best Local Similarity 63.6%; Pred. No. 3.2e+02;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      2 GGPGTTSNRD 12
Db      36 GGPGTNLNLD 46
|||||

RESULT 24
US-10-437-963-195589
; Sequence 195589, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 195589
; LENGTH: 152
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_91522C.1.pep
US-10-437-963-195589

Query Match          55.1%; Score 38; DB 16; Length 152;
Best Local Similarity 63.6%; Pred. No. 3.2e+02;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      2 GGPGTTSNRD 12
|||||

Db          13 GSSGNTSNRVD 23
|||||
RESULT 25
US-10-767-701-38825
; Sequence 38825, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 38825
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C74465_1.pep
US-10-767-701-38825

Query Match          55.1%; Score 38; DB 16; Length 213;
Best Local Similarity 58.3%; Pred. No. 4.6e+02;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY      2 GGPGTTSNRD 13
|||||
Db      35 GGPGITGLVDA 46
|||||

Search completed: September 24, 2004, 02:25:18
Job time : 79.3019 secs
```

This Page Blank (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 01:33:11 ; Search time 19.3774 Seconds
(without alignments)
64.534 Million cell updates/sec

Title: US-10-088-639A-2_COPY_226_238

Perfect score: 69

Sequence: 1 EGGPGTTSNRLDA 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database :

PIR 78:**

1: pir1:**

2: pir2:**

3: pir3:**

4: pir4:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	62.3	149	2	C87666 ribonuclease HI (i
2	40	58.0	249	2	T03011 dnaC protein homol
3	40	58.0	249	2	AG0618 probable DNA repli
4	40	58.0	344	2	S57901 estradiol 17beta-d
5	40	58.0	351	2	T36010 hypothetical prote
6	39	56.5	1073	2	T33764 hypothetical prote
7	39	56.5	1112	2	H95964 probable outer mem
8	39	56.5	3938	2	T42761 Bassoon protein-
9	38	55.1	293	2	E69174 succinate-CoA liga
10	38	55.1	293	2	F83618 conserved hypothet
11	38	55.1	312	2	AG2884 hypothetical prote
12	38	55.1	312	2	D97660 hypothetical prote
13	38	55.1	352	2	F70134 flagellar motor sw
14	38	55.1	667	2	F70682 probable membrane
15	38	55.1	1872	2	T24683 hypothetical prote
16	37	53.6	139	2	E75346 organic hydroperox
17	37	53.6	194	2	I57523 HSP90 - mouse (fra
18	37	53.6	222	2	A72643 hypothetical prote
19	37	53.6	229	2	T40148 hypothetical prote
20	37	53.6	290	2	T03552 maltose transport
21	37	53.6	353	2	T09665 peroxidase (EC 1.1
22	37	53.6	385	2	H86731 conserved hypothet
23	37	53.6	411	2	S45318 keratin 12 - rabbi
24	37	53.6	443	2	T21499 hypothetical prote
25	37	53.6	604	2	S25203 smR protein - Str
26	37	53.6	616	2	A72627 probable 2-oxoacid
27	36	52.2	143	2	A49438 p53 tumor suppress
28	36	52.2	148	2	A83418 ribonuclease H PA1
29	36	52.2	159	2	I49023 tumor suppressor p

hypothetical prote
hypothetical prote
Holliday junction
DNA-binding protei
hypothetical prote
synxain synt4 [lm
peroxidase (EC 1.1
probable methionin
carboxy beta-keto
ADK1 [imported] -
conserved hypothet
hypothetical prote
hypothetical prote
fibrinogen alpha c
chemotaxis protein
hypothetical prote
hydrogenase expre
hypothetical prote
hypothetical prote
fibrinogen alpha c
hypothetical prote
DNA-binding protei
hypothetical prote
hypothetical prote
hypothetical prote
hypothetical prote
major allergen Cor
major allergen Cor
major allergen Cor
major allergen Cor
desiccation stress
hypothetical prote
hypothetical prote
probable [imported]
probable multiple
hypothetical prote
ABC transporter, p
B-cell receptor, p
estradiol 17beta-d
transcription fact
transcription fact
transcription fact
transcription fact
hypothetical prote
avrRpt2-induced pr
aspartate ammonia-
histidine kinase h
NADP-reducing hydr
long-chain-fatty-a
long-chain-fatty-a
hypothetical prote
hypothetical prote
hypothetical prote
IAA4 [imported] -
hypothetical prote
transcription fact
myotubularin prote
hypothetical prote
protein kinase PKN
collagen alpha 2 c
BH-protocadherin-a
BH-protocadherin p
pyruvate carboxyla
BH-protocadherin p
profilaggrin relat
collagen alpha 1(I
collagen alpha 1'(

103 35 50.7 1520 2 AF3008 polyketide synthas
104 35 50.7 1520 2 G98275 hypothetical prote
105 34.5 50.0 1464 1 VEHY vimentin - golden
106 34.5 50.0 702 2 G01840 T-box protein 2 -
107 34.5 50.0 758 2 B83319 ATP-binding protei
108 34.5 50.0 1474 2 F69009 probable membrane
109 34 49.3 63 2 C95851 hypothetical prote
110 34 49.3 70 2 JH0128 E protein - Escher
111 34 49.3 129 2 AF1177 transcription regu
112 34 49.3 132 1 PSHOA phospholipase A2 (
113 34 49.3 148 2 I38881 caudal-type homeot
114 34 49.3 164 2 AF2293 hypothetical prote
115 34 49.3 184 2 S77928 exoskeletal protei
116 34 49.3 188 2 JC6547 high sulfur protei
117 34 49.3 203 2 I38868 caudal-type homeot
118 34 49.3 237 2 B86420 unknown protein, 1
119 34 49.3 259 2 T44461 transcription init
120 34 49.3 273 2 S05207 vimentin - pig (fr
121 34 49.3 284 2 S58650 vimentin - pig (fr
122 34 49.3 287 2 A69523 succinate-CoA liga
123 34 49.3 292 1 D55582 cytochrome-c oxida
124 34 49.3 300 2 B95971 probable sugar upt
125 34 49.3 306 2 T09067 extensin-like prot
126 34 49.3 316 2 T34838 probable transfera
127 34 49.3 333 2 G83391 conserved hypothet
128 34 49.3 335 2 T26517 hypothetical prote
129 34 49.3 355 2 JC4781 peroxidase (EC 1.1
130 34 49.3 358 2 JC4781 peroxidase (EC 1.1
131 34 49.3 358 2 H72337 oxidoreductase - T
132 34 49.3 376 2 G86483 hypothetical prote
133 34 49.3 377 2 E82402 conserved hypothet
134 34 49.3 378 2 C64142 conserved hypothet
135 34 49.3 382 2 A70071 conserved hypothet
136 34 49.3 384 2 AE3189 aminopeptidase Atu
137 34 49.3 389 2 I49640 transcription fact
138 34 49.3 399 2 JC4592 transcription fact
139 34 49.3 401 2 D83618 beta-ketoadipyl Co
140 34 49.3 403 2 S02709 ear-2 protein - hu
141 34 49.3 424 2 T07366 probable phosphati
142 34 49.3 425 2 E64403 translation releas
143 34 49.3 426 2 H84846 probable polygalac
144 34 49.3 426 2 A39695 transforming protei
145 34 49.3 427 2 F87465 GTP-binding protei
146 34 49.3 443 2 B87706 conserved hypothet
147 34 49.3 450 2 B87326 hypothetical prote
148 34 49.3 466 2 A25074 vimentin - human
149 34 49.3 481 2 AH3350 alkaline phosphata
150 34 49.3 483 2 A55033 keratin 12 - mouse

ALIGNMENTS

RESULT 1
C87666
ribonuclease HI [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
C:Accession: C87666
R:NIERMAN, W.C.; FELDBLYUM, T.V.; PAULSEN, I.T.; NELSON, K.E.; EISEN, J.; HEIDELBERG, J.
B.; LAUB, M.T.; DEBOY, R.T.; DODSON, R.J.; DURKIN, A.S.; GWINN, M.L.; HAFT, D.H.; KOLON
N, J.; ERMOLOVA, M.; WHITE, O.; SALZBERG, S.L.; SHAPIRO, L.; VENTER, J.C.; FRASER, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: C87666
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-149 <STO>
A:Cross-references: GB:AE005673; NID:g13425071; PIDN:AAK25327.1; GSPDB:GN00148
C:Genetics: CC3365
C:Superfamily: ribonuclease H

Query Match 62.3%; Score 43; DB 2; Length 149;
Best Local Similarity 63.6%; Pred. No. 3.7;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 2 GGPGTTSNRLD 12
DB 38 GEPGTNNRME 48
|||::|::|
RESULT 2
T03011
dnaC protein homolog - Salmonella typhimurium
C:Species: Salmonella typhimurium
C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 26-Aug-1999
C:Accession: T03011
R:FIGUEROA-BOSSI, N.; BOSSI, L.
submitted to the EMBL Data Library, June 1998
A:Reference number: Z14818
A:Accession: T03011
A>Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-249 <FIG>
A:Cross-references: EMBL:AF001386; NID:g3294471; PIDN:AAK26072.1; PID:g32944483
C:Superfamily: DNA replication protein dnaC

Query Match 58.0%; Score 40; DB 2; Length 249;
Best Local Similarity 66.7%; Pred. No. 21;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 GGPGTTSNRLDA 13
DB 109 GGPGTGKNHLAA 120
|||::|::|
RESULT 3
AG0618
probable DNA replication protein STY1023 [imported] - Salmonella enterica subsp. enterica
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AG0618
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AG0618
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-249 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD05417.1; PID:g16502178; GSPDB:GN00176
C:Genetics: STY1023
C:Superfamily: DNA replication protein dnaC

Query Match 58.0%; Score 40; DB 2; Length 249;
Best Local Similarity 66.7%; Pred. No. 21;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 GGPGTTSNRLDA 13
DB 109 GGPGTGKNHLAA 120
|||::|::|
RESULT 4
S57901
estradiol 17beta-dehydrogenase (EC 1.1.1.62) - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 24-Nov-2003
C:Accession: S57901

R;Chersevich, S.; Nokelainen, P.; Poutanen, M.; Orava, M.; Autio-Harmainen, H.; Rajanien
Endocrinology 135, 1477-1487, 1994
A;Title: Rat 17-beta-hydroxysteroid dehydrogenase type 1: primary structure and regulation
A;Reference number: S57901; MUID:95009707; PMID:7925110
A;Accession: S57901
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-344 <GHE>
A;Cross-references: EMBL:X78811; NID:G561532; PIDN:CAA55389.1; PID:G561533
C;Superfamily: 17beta-dehydrogenase; short-chain alcohol dehydrogenase homology
C;Keywords: oxidoreductase
F;4-189/Domain: short-chain alcohol dehydrogenase homology <SADH>

Query Match 58.0%; Score 40; DB 2; Length 344;
Best Local Similarity 61.5%; Pred. No. 29;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 EGGPGTTSNRDLA 13
Db 198 EGGPGGALERADA 210
|||||
|

RESULT 5
T36010
hypothetical protein SCC22.20 - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C;Accession: T36010
R;Seeger, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, July 1999
A;Reference number: Z21574
A;Accession: T36010
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-351 <SEE>
A;Cross-references: EMBL:AL096839; PIDN:CAB50763.1; GSPDB:GN00070; SCOEDB:SCC22.20
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCOEDB:SCC22.20

Query Match 58.0%; Score 40; DB 2; Length 351;
Best Local Similarity 63.6%; Pred. No. 30;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 GGPGLTAVRMD 243
Db 233 GGPGLTAVRMD 243
|||||
|

RESULT 6
T33764
hypothetical protein C02E11.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C;Accession: T33764
R;Pauley, A.; Harper, M.
submitted to the EMBL Data Library, October 1998
A;Description: The sequence of C. elegans cosmid C02E11.
A;Reference number: Z21403
A;Accession: T33764
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1073 <PAU>
A;Cross-references: EMBL:AF101304; PIDN:AAC69200.1; GSPDB:GN00023; CESP:C02E11.1
A;Experimental source: strain Bristol N2; clone C02E11
C;Genetics:
A;Gene: CESP:C02E11.1
A;Map position: 5
A;Introns: 37/1; 174/3; 248/3; 511/3; 702/2; 1043/3

Query Match 56.5%; Score 39; DB 2; Length 1073;
Best Local Similarity 80.0%; Pred. No. 1.4e+02;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 GPGTTSNRRLD 12
Db 458 GPGTYSARLD 467
|||||
|

RESULT 7

H95964
probable outer membrane secretion protein SMB21543 [imported] - Sinorhizobium meliloti (S.
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C;Accession: H95964
R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A;Title: The complete sequence of the 1.683-kb pSymb megaplasmid from the N2-fixing endo
A;Reference number: A95842; MUID:21396508; PMID:11481431
A;Accession: H95964
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1112 <KUR>
A;Cross-references: GB:AL591985; PIDN:CAC49384.1; PID:G15140870; GSPDB:GN00167
A;Experimental source: strain 1021, megaplasmid pSymb
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kise, E.; Komp, C.; Lelaure,
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: SMB21543
A;Genome: plasmid

Query Match 56.5%; Score 39; DB 2; Length 1112;
Best Local Similarity 77.8%; Pred. No. 1.5e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 GPGGTTSNR 10
Db 913 GPGGTTSNR 921
|||||
|

RESULT 8

T42761
Bassoon protein - rat
N;Alternate names: brain-specific synapse-associated protein
C;Species: Rattus norvegicus (Norway rat)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 21-Jul-2000
C;Accession: T42761
R;Dieck, S.; Sammarti-Vila, L.; Langnaese, K.; Richter, K.; Kindler, S.; Soyke, A.; Wex,
J. Cell Biol. 142, 499-509, 1998
A;Title: Bassoon, a novel zinc-finger CAG/Glutamine-repeat protein selectively localized
A;Reference number: Z22249; MUID:98345363; PMID:9679147
A;Accession: T42761
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-3938 <DIE>
A;Cross-references: EMBL:Y16563; NID:G3413503; PIDN:CAA76287.1; PID:G3413504
A;Experimental source: strain Sprague Dawley; brain
C;Function:
A;Description: may be involved in cytomatrix organization at the site of neurotransmitter
A;Note: component of the presynaptic cytoskeleton
C;Keywords: coiled coil; zinc finger

Query Match 56.5%; Score 39; DB 2; Length 3938;
Best Local Similarity 70.0%; Pred. No. 5.6e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 GPGTTSNRRLD 12
Db 69 GPGSVSRRLD 78
|||||
|

```

erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClellan, E. W.
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, E. W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AG2884
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-312 <KUR>
A:Cross-references: GB:AE008688; PIDN:AAL43493.1; PID:g17741000; GSPDB:GN00186
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atu2506
A:Map position: circular chromosome
C:Superfamily: inner membrane protein upgA

Query Match 55.1%; Score 38; DB 2; Length 312;
Best Local Similarity 50.0%; Pred. No. 59;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 EGGPGTTSNRDL 12
Db 252 QGGPGTASSETIN 263

RESULT 12
D97660
hypothetical protein AGR_C4553 [imported] - Agrobacterium tumefaciens (strain C58, Cere
C:Species: Agrobacterium_tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002
C:Accession: D97660
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorillo, B.; Goldman,
; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: D97660
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-312 <KUR>
A:Cross-references: GB:AE007869; PIDN:AAK8237.1; PID:g15157693; GSPDB:GN00169
C:Genetics:
A:Gene: AGR_C_4553
A:Map position: circular chromosome
C:Superfamily: inner membrane protein upgA

Query Match 55.1%; Score 38; DB 2; Length 312;
Best Local Similarity 50.0%; Pred. No. 59;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 EGGPGTTSNRDL 12
Db 252 QGGPGTASSETIN 263

RESULT 13
F70134
flagellar motor switch protein (flM) homolog - Lyme disease spirochete
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 20-Aug-1999
C:Accession: F70134
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White,
son, D.; Peterson, J.; Kexlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.
A:Reference number: A70100; MUID:98065943; PMID:9403685
A:Accession: F70134
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA

```

```

succinate-CoA ligase (ADP-forming) (EC 6.2.1.5) alpha chain - Methanobacterium thermoaut
C:Species: Methanobacterium thermoautotrophicum
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 03-Jun-2002
C:Accession: E69174
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiواني, N.
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct
A:Reference number: A69000; MUID:98037514; PMID:9371463
A:Accession: E69174
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-293 <MTH>
A:Cross-references: GB:AE000839; GB:AE000666; NID:g2621637; PIDN:AAB85069.1; PID:g262164
A:Experimental source: strain Delta H
C:Genetics:
A:Note: MTH563
C:Superfamily: succinate-CoA ligase (ADP-forming) alpha chain
C:Keywords: acid-thiol ligase; ATP; coenzyme A; phosphohistidine; phosphoprotein; tricar
F;249/Active site: His (phosphohistidine intermediate) #status predicted

Query Match 55.1%; Score 38; DB 2; Length 293;
Best Local Similarity 53.8%; Pred. No. 55;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 EGGPGTTSNRDLA 13
Db 255 EGGTGASSKREA 267

RESULT 10
F83618
conserved hypothetical protein PA0209 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: F83618
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: F83618
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-293 <STO>
A:Cross-references: GB:AE004459; GB:AE004091; NID:g9946043; PIDN:AAG03598.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA0209
C:Superfamily: citG protein

Query Match 55.1%; Score 38; DB 2; Length 293;
Best Local Similarity 61.5%; Pred. No. 55;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 EGGPGTTSNRDLA 13
Db 193 EGGAGEQNARLDA 205

RESULT 11
AG2884
hypothetical protein Atu2506 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
C:Accession: AG2884
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.

```

A:Residues: 1-352 <KLE>
A:Cross-references: GB:AE001137; GB:AE000783; NID:g2688160; PIDN:AAC66670.1; PID:g268818
A:Experimental source: strain B31
C:Superfamily: flagellar motor switch protein flm

Query Match 55.1%; Score 38; DB 2; Length 352;
Best Local Similarity 72.7%; Pred. No. 67;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 GPGTTSNRLDA 13
DB 250 GVGTTSENLD 260
:|||||:|

RESULT 14
F70682
Probable membrane protein - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: F70682
R: Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: F70682
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-667 <COL>
A:Cross-references: GB:Z81368; GB:AL123456; NID:g3261656; PIDN:CAB03731.1; PID:g1655665
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: Kv2395
C:Superfamily: Pyrococcus horikoshii hypothetical protein PH0361

Query Match 55.1%; Score 38; DB 2; Length 667;
Best Local Similarity 53.8%; Pred. No. 1.3e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 EGGPGTTSNRLDA 13
DB 475 QGAPGATANALAA 487
:|||||:|

RESULT 15
T24683
Hypothetical protein T08A11.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
C:Accession: T24683; T26904
R: Chui, C.; Sulston, J.
A:Reference number: Z19922
A:Accession: T24683
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-1872 <WIL>
A:Cross-references: EMBL:Z50875; PIDN:CAA90776.1; GSPDB:GN00021; CESP:T08A11.1
A:Experimental source: clone T08A11
R:McMurray, A.
A:Reference number: Z20284
A:Accession: T26904
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-1872 <W12>
A:Cross-references: EMBL:AL021180; PIDN:CAA15982.1; GSPDB:GN00021; CESP:T08A11.1
A:Experimental source: clone Y44F5A
C:Genetics:
A:Gene: CESP:T08A11.1
A:Map position: 3

A:Introns: 58/3; 330/3; 366/3; 498/3; 544/3; 605/3; 896/3; 924/1; 1023/1; 1139/2; 1217/3;
Query Match 55.1%; Score 38; DB 2; Length 1872;
Best Local Similarity 50.0%; Pred. No. 3.8e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 EGGPGTTSNRLD 12
DB 338 KGAPGLITNRID 349
:|||||:|

RESULT 16
E75346
Organic hydroperoxide resistance protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C:Accession: E75346
R: White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; Lam, M.; Shan, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Mat S.; Smith, H.O.; Venter, J.C.; Fraser, C.M. Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: E75346
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-139 <WHI>
A:Cross-references: GB:AE002025; GB:AE000513; NID:g6459627; PIDN:AAF11408.1; PID:g645963
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR1857
A:Map position: 1
C:Superfamily: hypothetical protein yk1A

Query Match 53.6%; Score 37; DB 2; Length 139;
Best Local Similarity 53.8%; Pred. No. 38;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 EGGPGTTSNRLDA 13
DB 39 DGGPGTNPEQLFA 51
:|||||:|

RESULT 17
I57523
HSP90 - mouse (fragment)
C:Species: Mus sp. (mouse)
C>Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 20-Aug-1999
C:Accession: I57523
R:Shaknovich, R.; Shue, G.; Kohtz, D.S. Mol. Cell. Biol. 12, 5059-5068, 1992
A:Title: Conformational activation of a basic helix-loop-helix protein (MyoD1) by the C-terminus of the myosin heavy chain.
A:Reference number: I57523; MUID:93024452; PMID:1406681
A:Accession: I57523
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-194 <RES>
A:Cross-references: GB:S46109; NID:g257730; PIDN:AAB23704.1; PID:g2577731
C:Superfamily: heat shock protein 90

Query Match 53.6%; Score 37; DB 2; Length 194;
Best Local Similarity 75.0%; Pred. No. 53;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 GGGPGTTSN 9
DB 1 GGGPGVTN 8
:|||||:|

RESULT 18
A72643
Hypothetical protein APE0577 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix

C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999

C:Accession: A72643
R:Kawarabayashi, Y.; Hino, Y.; Hoikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Maeda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K DNA Res. 6, 83-101, 1999

A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum

A:Reference number: A72450; MUID:99310339; PMID:10382966

A:Accession: A72643
A:Status: preliminary

A:Molecule type: DNA
A:Residues: 1-222 <KAW>

A:Cross-references: DDBJ:AP000060; NID:g5104188; PIDN:BAA79545.1; PID:d1043331; PID:g510

A:Experimental source: strain K1

C:Genetics:
A:Gene: APE0577

Query Match 53.6%; Score 37; DB 2; Length 222;
Best Local Similarity 63.6%; Pred. No. 62;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 GPGGTTSNRLDA 13
||| :|||
DB 174 GPGGTLASRLTA 184

RESULT 19

T40148
hypothetical protein SPBC2G2.09c - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

C:Accession: T40148

R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Devlin, K.; Churcher, C.M.

submitted to the EMBL Data Library, March 1998

A:Reference number: Z21842

A:Accession: T40148

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-229 <WOO>

A:Cross-references: EMBL:AL022103; PIDN:CAA17889.1; GSPDB:GN00067; SPDB:SPBC2G2.09c

A:Experimental source: strain 972h; cosmid c2G2

C:Genetics:

A:Gene: SPDB:SPBC2G2.09c

A:Map position: 2

A:Introns: 96/3; 150/1; 168/3

Query Match 53.6%; Score 37; DB 2; Length 229;
Best Local Similarity 70.0%; Pred. No. 64;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 GPGGTTSNRLD 12
||| :|||
DB 29 GPLTTSNEID 38

RESULT 20

T03552

maltoase transport inner membrane protein - Rhodobacter capsulatus

C:Species: Rhodobacter capsulatus

C>Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 08-Oct-1999

C:Accession: T03552

R:VLcek, C.; Paces, V.; Maltsev, N.; Paces, J.; Haselkorn, R.; Fonstein, M.

Proc. Natl. Acad. Sci. U.S.A. 94, 9384-9388, 1997

A:Title: Sequence of a 189-kb segment of the chromosome of Rhodobacter capsulatus SB1003

A:Reference number: Z14955; MUID:97404404; PMID:9256491

A:Accession: T03552

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-290 <VLC>

A:Cross-references: EMBL:AF010496; NID:g3128256; PIDN:AAC16205.1; PID:g3128353

C:Genetics:

A:Map position: 1

C:Superfamily: inner membrane protein ugpa

Query Match 53.6%; Score 37; DB 2; Length 290;
Best Local Similarity 70.0%; Pred. No. 81;

Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 GPGGTTSNRL 11
||| :|||
DB 236 GPGGTASTTL 245

RESULT 21

T09665

peroxidase (BC 1.11.1.7) pxdc precursor - alfalfa

C:Species: Medicago sativa (alfalfa)

C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 26-Aug-1999

C:Accession: T09665

R:Abrahams, S.L.; Hayes, C.M.; Watson, J.M.

submitted to the EMBL Data Library, September 1994

A:Description: Organ-specific expression of three peroxidase-encoding cDNAs from lucerne

A:Reference number: Z16809

A:Accession: T09665

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: mRNA

A:Residues: 1-353 <ABR>

A:Cross-references: EMBL:L36157; NID:g537316; PID:g537317

A:Experimental source: subspecies sativa; cultivar Siriver

C:Genetics:

A:Gene: pxdc

C:Superfamily: peroxidase

C:Keywords: oxidoreductase

F:1-28/Domain: signal sequence #status predicted <SIG>

F:29-353/Product: peroxidase pxdc #status predicted <MAT>

Query Match 53.6%; Score 37; DB 2; Length 353;
Best Local Similarity 63.6%; Pred. No. 16+02;

Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 GPGGTTSNRLD 12
||| :|||
DB 239 GPGGTLTDLTD 249

RESULT 22

H86731

conserved hypothetical protein yihF [imported] - Lactococcus lactis subsp. lactis (strain

C:Species: Lactococcus lactis subsp. lactis

C>Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001

C:Accession: H86731

R:Boletín, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarre, K.; Weissbach, J.; Ehrlich

Genome Res. 11, 731-753, 2001

A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ssi

A:Reference number: A86625; MUID:21235186; PMID:11337471

A:Accession: H86731

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-385 <STO>

A:Cross-references: GB:AE005176; PID:g12723784; PIDN:AAK04954.1; GSPDB:GN00146

A:Experimental source: strain ILI403

C:Genetics:

A:Gene: yihF

C:Superfamily: yhad protein

Query Match 53.6%; Score 37; DB 2; Length 385;
Best Local Similarity 53.8%; Pred. No. 1.1e+02;

Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 EGGPGTTSNRLDA 13
:|||||:
DB 45 DGGEGTIDALIDA 57

RESULT 23

S45318

keratin 12 - rabbit (fragment)

C:Species: Oryctolagus cuniculus (domestic rabbit)
C>Date: 06-Jan-1995 #sequence_revision 17-Jul-1998 #text_change 10-Dec-1999
C:Accession: S45318
R:Wu, R.L.; Zhu, G.; Galvin, S.; Xu, C.; Haseeba, T.; Chaloin-Dufau, C.; Dhoulailly, D.; M
Differentiation 55, 137-144, 1994
A:Title: Lineage-specific and differentiation-dependent expression of K12 keratin in rat
A:Reference number: S45318; MUID:94192891; PMID:7511548
A:Accession: S45318
A:Molecule type: mRNA
A:Residues: 1-411 <WUR>
A:Cross-references: EMBL:X77665; NID:g495260; PIDN:CAA54741.1; PID:g495261
C:Superfamily: Cytoskeletal keratin
C:Keywords: coiled coil; intermediate filament

Query Match 53.6%; Score 37; DB 2; Length 411;
Best Local Similarity 58.3%; Pred. No. 1.2e+02;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 GPGGTTSNRLDA 13
|||||
Db 196 GGPGEVSEMDA 207

RESULT 24

T21499
hypothetical protein F28D1.8 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T21499
R:Baynes, C.
submitted to the EMBL Data Library, April 1996
A:Reference number: Z19430
A:Accession: T21499
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-443 <WIL>
A:Cross-references: EMBL:Z70684; PIDN:CAA94603.1; GSPDB:GN00022; CESP:F28D1.8
A:Experimental source: clone F28D1
C:Genetics:
A:Gene: CESP:F28D1.8
A:Map position: 4
A:Introns: 71/1; 103/3; 162/3; 215/2; 360/2

Query Match 53.6%; Score 37; DB 2; Length 443;
Best Local Similarity 70.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 GPGGTTSNRLD 12
|||||
Db 297 GPGTTSKSTD 306

RESULT 25

S25203
srnR protein - Streptomyces ambofaciens
C:Species: Streptomyces ambofaciens
C>Date: 28-May-1993 #sequence_revision 28-May-1993 #text_change 22-Oct-1999
C:Accession: S25203; S21598
R:Geistlich, M.; Losick, R.; Turner, J.R.; Rao, R.N.
Mol. Microbiol. 6, 2019-2029, 1992
A:Title: Characterization of a novel regulatory gene governing the expression of a polyk
A:Reference number: S25202; MUID:92374852; PMID:1508047
A:Accession: S25203
A:Molecule type: DNA
A:Residues: 1-604 <GEI>
A:Cross-references: EMBL:X63451; NID:g46699; PIDN:CAA45051.1; PID:g46701
C:Genetics:
A:Gene: srnR

Query Match 53.6%; Score 37; DB 2; Length 604;
Best Local Similarity 72.7%; Pred. No. 1.8e+02;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 GPGGTTSNRLD 12
|||||
Db 385 GGPGRGTSPELD 395

Search completed: September 24, 2004, 01:54:38
Job time : 23.3774 secs

This Page Blank (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2005, 17:57:45 ; Search time 97.9062 Seconds
(without alignments)
51.354 Million cell updates/sec

Title: US-10-088-639A-2_COPY_226_238

Perfect score: 69

Sequence: 1 EGGPGTTSNRLDA 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

1: Geneseqp1980s:*

2: Geneseqp1990s:*

3: Geneseqp2000s:*

4: Geneseqp2001s:*

5: Geneseqp2002s:*

6: Geneseqp2003as:*

7: Geneseqp2003bs:*

8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	69	100.0	249	AAB68087	Aab68087 An anti-a
2	44	63.8	527	AAU64492	Aau64492 Propionib
3	44	63.8	527	ABM64701	Abm64701 Propionib
4	44	63.8	527	ABM61011	Abm61011 Propionib
5	44	63.8	3149	ADJ70520	Adj70520 Human hea
6	43	62.3	749	ABB60469	Abb60469 Drosophil
7	41	59.4	166	ADG22481	Adg22481 Cyanophag
8	41	59.4	305	ABO74099	AbO74099 Pseudomon
9	41	59.4	398	ADN22235	Adn22235 Bacterial
10	41	59.4	400	ADN24993	Adn24993 Bacterial
11	41	59.4	19938	ABP76682	Abp76682 Streptomy
12	40	58.0	92	ADG22495	Adg22495 Cyanophag
13	40	58.0	224	AAy74086	Aay74086 Human pro
14	39	56.5	77	AAm85724	Aam85724 Human imm
15	39	56.5	130	ABU41557	Abu41557 Protein e
16	39	56.5	219	AAU83139	Aau83139 Novel sec
17	39	56.5	296	ABP43778	Abp43778 FLJ14840
18	39	56.5	302	ABB68531	Abb68531 Drosophil
19	39	56.5	351	ABO68553	AbO68553 Pseudomon
20	39	56.5	396	AAAB94593	AaB94593 Human pro
21	39	56.5	396	AAAG67257	Aag67257 Amino aci
22	39	56.5	454	AAU35107	Aau35107 Enterococ
23	39	56.5	454	ABU29209	Abu29209 Protein e
24	39	56.5	461	ADH87666	Adh87666 Enterococ
25	39	56.5	600	AAAB85503	Aab85503 Human pro

ALIGNMENTS

RESULT 1

AAB68087

ID AAB68087 standard; protein; 249 AA.

XX AC AAB68087;

XX DT 09-JUL-2001 (first entry)

XX DE An anti-alpha6beta4 integrin light chain linked to a heavy chain.

XX KW Gastrointestinal epithelial tumour cell; alpha6beta4 integrin;

XX KW tumour-associated antigen; metastatic disease; malignant disease;

XX KW tumour typing; tumour screening; tumour.

XX OS Synthetic.

XX OS Macaca fascicularis.

XX FH Key Location/Qualifiers

FT Region 23..33 /note= "Complementarity determining region (CDR) 1 of the

FT light chain"

FT Region 49..55 /note= "Complementarity determining region (CDR) 2 of the

FT light chain"

FT Region 88..98 /note= "Complementarity determining region (CDR) 3 of the

FT light chain"

FT Peptide 110..127 /note= "linker"

FT Region 158..162 /note= "Complementarity determining region (CDR) 1 of the

FT heavy chain"

FT Region 177..193 /note= "Complementarity determining region (CDR) 1 of the

FT heavy chain"

FT Region 226..238 /note= "Complementarity determining region (CDR) 1 of the

FT heavy chain"

XX WO200130854-A2.

XX 03-MAY-2001.

XX 26-OCT-2000; 2000WO-SE002082.

XX 28-OCT-1999; 99SE-00003895.

XX (ACTI-) ACTIVE BIOTECH AB.

```

XX Brodin TN, Karlstroem PJ, Ohlsson LG, Tordsson MJ, Kearney PP;
PI Nilsson BHK;
XX
XX WPI; 2001-308619/32.
DR N-PSDB; AAP84797.
XX
XX Novel antibody for diagnosis, treatment of human metastatic and malignant
PT diseases, has binding structure for target structure displayed on cell
PT surface of human gastrointestinal epithelial tumor and normal cells.
XX
XX Claim 1; Page 55-56; 75pp; English.
XX
XX The present sequence represents a Monkey antibody light chain linked to a
CC heavy chain. The antibody binds to a target structure displayed in and on
CC the cell surface of human gastrointestinal epithelial tumour cells and in
CC a subpopulation of normal human gastrointestinal epithelial cells. The
CC target structure especially comprises alpha6beta4 integrin. This is a
CC tumour-associated antigen. The antibody, and its fragments, are useful
CC for treating conditions based on anti-angiogenic mechanism in humans.
CC They are useful for treating human metastatic and malignant disease, for
CC in vitro, in vivo diagnosis and prognosis of human malignant disease,
CC comprising tumour typing, tumour screening, tumour diagnosis and
CC prognosis and monitoring premalignant conditions. Quantitative in vivo
CC diagnosis is carried out by determining the localization of antibody to
CC tumour deposits in humans
XX
XX Sequence 249 AA;
SQ
Query Match 100.0%; Score 69; DB 4; Length 249;
Best Local Similarity 100.0%; Pred. No. 0.0041;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EGGPGTTSNRLDA 13
DB 226 EGGPGTTSNRLDA 238
|||||
XX
RESULT 2
AAU64492
ID AAU64492 standard; protein; 527 AA.
XX
AC AAU64492;
XX
XX 27-FEB-2002 (first entry)
DT
DE
DE Propionibacterium acnes immunogenic protein #25388.
XX
XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
XX Propionibacterium acnes.
OS
XX WO200181581-A2.
PN
XX 01-NOV-2001.
PD
XX
XX 20-APR-2001; 2001WO-US012865.
PF
XX
XX 21-APR-2000; 2000US-0199047P.
PR
XX 02-JUN-2000; 2000US-0208841P.
PR
XX 07-JUL-2000; 2000US-0216747P.
PR
XX
XX (CORI-) CORIXA CORP.
PA
XX
XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'Maisonneuve J, Zhang Y, Jen S, Carter D;
XX
XX WPI; 2001-616774/71.
DR
DR N-PSDB; AAS59645.
XX

```

```

PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.
XX
XX Example 1; SEQ ID NO 25687; 1069pp; English.
XX
XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 527 AA;
SQ
Query Match 63.8%; Score 44; DB 4; Length 527;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 2 GGGPGTTSNRL 11
DB 7 GGGPGTASHEL 16
|||||
XX
RESULT 3
ABM64701
ID ABM64701 standard; protein; 527 AA.
XX
AC ABM64701;
XX
XX 20-OCT-2003 (first entry)
DT
DE
DE Propionibacterium acnes immunogenic polypeptide #29377.
XX
XX Acne vulgaris; antisporrheic; dermatological; antibacterial;
KW immunostimulant; immune response; vaccine; immunogenic.
XX
XX Propionibacterium acnes.
OS
XX WO2003033515-A1.
PN
XX 24-APR-2003.
PD
XX
XX 11-OCT-2002; 2002WO-US032727.
PF
XX
XX 15-OCT-2001; 2001US-00978825.
PR
XX
XX (CORI-) CORIXA CORP.
PA
XX Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
PI Zhang Y, Wang S, Jen S, Lodes WJ, Benson DR, Jones R, Carter D;
PI Barth B, Vallieve-Douglas J;
XX
XX WPI; 2003-381789/36.
DR
XX
XX New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
PT or for stimulating an immune response specific for a P. acnes protein.
XX

```

PS Claim 7; SEQ ID NO 29377; 1481pp; English.

XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733) encoding a Propionibacterium acnes protein. The invention also relates to CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to CC immunogenic fragments of P. acnes polypeptides. The invention CC additionally encompasses expression vectors and host cells comprising a CC polynucleotide of the invention; antibodies against polypeptides of the CC invention; fusion proteins comprising a polypeptide of the invention; a CC method for stimulating an immune response specific for a P. acnes CC polypeptide and an isolated T cell population comprising T cells prepared CC via this method; a vaccine composition (comprising P. acnes polypeptides, CC polynucleotides, antibodies, fusion proteins, T cell populations, or CC antigen-presenting cells that express the polypeptide); a method and kit CC for detecting or determining the presence or absence of P. acnes in a CC patient; and a method for inhibiting the development of P. acnes in a CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion CC proteins, T cell populations or antigen-presenting cells that express the CC polypeptides are useful for diagnosing, preventing or treating acne, CC stimulation of an immune response against P. acnes, or for treating acne, CC and the kit is useful for performing a diagnostic assay. The present CC sequence represents a specifically claimed P. acnes polypeptide which is CC thought to contain an immunogenic region. Note: The sequence data for CC this patent did not form part of the printed specification, but was CC obtained in electronic format directly from WIPO at CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 527 AA;

Query Match 63.8%; Score 44; DB 6; Length 527;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 GGGGTTSNRL 11
DB 7 GGGGTASHRL 16
||||| :|||

RESULT 4
ID ABM61011 standard; protein; 527 AA.
AC ABM61011;
XX 20-OCT-2003 (first entry)
XX Propionibacterium acnes predicted ORF-encoded polypeptide #25687.
XX Acne vulgaris; antisemborrhoeic; dermatological; antibacterial;
KW immunostimulant; immune response; vaccine.
XX Propionibacterium acnes.
OS WO2003033515-A1.
PN 24-APR-2003.
PD 11-OCT-2002; 2002WO-US032727.
XX 15-OCT-2001; 2001US-00978825.
PR (CORI-) CORIXA CORP.
XX Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
PI Barth B, Valliave-Douglas J;
XX WPI; 2003-381789/36.
DR N-PSDB; ACF64574.

PT New Propionibacterium acnes polypeptides and polynucleotides encoding the PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris, PT or for stimulating an immune response specific for a P. acnes protein.

XX Example 1; SEQ ID NO 25687; 1481pp; English.

PS The invention relates to an isolated polynucleotide (ACF64435-ACF64733) encoding a Propionibacterium acnes protein. The invention also relates to CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to CC immunogenic fragments of P. acnes polypeptides. The invention CC additionally encompasses expression vectors and host cells comprising a CC polynucleotide of the invention; antibodies against polypeptides of the CC invention; fusion proteins comprising a polypeptide of the invention; a CC method for stimulating an immune response specific for a P. acnes CC polypeptide and an isolated T cell population comprising T cells prepared CC via this method; a vaccine composition (comprising P. acnes polypeptides, CC polynucleotides, antibodies, fusion proteins, T cell populations, or CC antigen-presenting cells that express the polypeptide); a method and kit CC for detecting or determining the presence or absence of P. acnes in a CC patient; and a method for inhibiting the development of P. acnes in a CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion CC proteins, T cell populations or antigen-presenting cells that express the CC polypeptides are useful for diagnosing, preventing or treating acne, CC stimulation of an immune response against P. acnes, or for treating acne, CC and the kit is useful for performing a diagnostic assay. The present CC sequence represents a polypeptide predicted to be encoded by an ORF (open CC reading frame) contained within the P. acnes polynucleotides of the CC invention. Note: The sequence data for this patent did not form part of CC the printed specification, but was obtained in electronic format directly CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 527 AA;

Query Match 63.8%; Score 44; DB 6; Length 527;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 GGGGTTSNRL 11
DB 7 GGGGTASHRL 16
||||| :|||

RESULT 5
ADJ70520
ID ADJ70520 standard; protein; 3149 AA.
AC ADJ70520;
XX 06-MAY-2004 (first entry)
DT Human heat mitochondrial protein as a therapeutic target SeqID2326.
DE mitochondrial; human; screening assay; diabetes mellitus;
KW Huntington's disease; osteoarthritis;
KW Leber's hereditary optic neuropathy; LHON;
KW mitochondrial encephalopathy lactic acidosis and stroke; MELAS;
KW myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;
KW neuroprotective; nootropic; antidiabetic; anticonvulsant; antiarthritic;
KW osteopathic; ophthalmological; cytostatic.
XX Homo sapiens.
OS WO2003087768-A2.
PN 23-OCT-2003.
XX 04-APR-2003; 2003WO-US010870.
XX 12-APR-2002; 2002US-0372843P.
PR 17-JUN-2002; 2002US-0389987P.

```

PR 20-SEP-2002; 2002US-0412418P.
XX
PA (MITO-) MITOKOR.
PA (BUCK-) BUCK INST AGE RES.
XX
PI Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;
PI Warnock DE;
XX
DR WPI; 2003-845369/78.
XX
XX Identifying a mitochondrial target for drug screening assays and for
PT treating diseases associated with altered mitochondrial function.
PT comprises detecting a modified polypeptide in a sample and correlating
PT with the disease.
XX
XX Claim 1; SEQ ID NO 2326; 180pp; English.
XX
XX This invention relates to novel mitochondrial targets that can be used
CC for therapeutic intervention in treating a disease associated with
CC altered mitochondrial function. Specifically, it refers to a method for
CC identifying proteins of the human heart mitochondrial proteome that are
CC useful for drug screening assays, as well as therapeutic targets. The
CC present invention describes a method for identifying such proteins that
CC can be used in the treatment of various diseases associated with altered
CC mitochondrial function including diabetes mellitus, Huntington's disease,
CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial
CC encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy
CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these
CC compositions have neuroprotective, neurotropic, antidiabetic,
CC anticonvulsant, antiarthritic, osteopathic, ophthalmological and
CC cytostatic activities. This polypeptide sequence is a human heart
CC mitochondrial protein of the invention.
XX
XX Sequence 3149 AA;
SQ
    Query Match      63.8%; Score 44; DB 7; Length 3149;
    Best Local Similarity 80.0%; Pred. No. 7.7e+02;
    Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 GPGTTSNRRLD 12
Db      |||:|||||
      72 GPGTTSNRRLD 81

RESULT 6
AB60469
ID ABB60469 standard; protein; 749 AA.
XX
AC ABB60469;
XX
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster polypeptide SEQ ID NO 8199.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical.
XX
XX Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US009231.
XX
XX 23-MAR-2000; 2000US-0191637P.
XX
XX 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
XX
DR

DR N-PSDB; ABL04572.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions.
XX
XX Disclosure; SEQ ID NO 8199; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 749 AA;
SQ
    Query Match      62.3%; Score 43; DB 4; Length 749;
    Best Local Similarity 66.7%; Pred. No. 2.5e+02;
    Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 GPGTTSNRRLDA 13
Db      |||||:|
      127 GPGTASNRLLQS 138

RESULT 7
ADG22481
ID ADG22481 standard; protein; 166 AA.
XX
AC ADG22481;
XX
XX 26-FEB-2004 (first entry)
XX
XX Cyanophaga S-2L encoded protein #226.
XX
XX genome; cyanophaga; 2; 6-diaminopurine; chemotherapy; AIDS.
XX
XX Cyanophaga S-2L.
XX
XX FR2839079-A1.
XX
XX 31-OCT-2003.
XX
XX 30-APR-2002; 2002FR-00005424.
XX
XX 30-APR-2002; 2002FR-00005424.
XX
XX (INSP ) INST PASTEUR.
XX
XX (CNFS ) CNFS CENT NAT RECH SCI.
XX
XX (GENO-) GENOSCOPE CENT NAT SEQUENCAGE GRP INTERE.
XX
XX
XX Marliere P, Kaminski PA, Galisson F, Bouzon M, Pochet S;
XX Weissenbach J, Saurin W, Robert C, Vico V;
XX
XX WPI; 2004-045746/05.
XX
XX N-PSDB; ADG22255.
XX
XX New genomic sequence for cyanophaga S-2L, useful for identifying genes
PT for synthesis of 2,6-diaminopurine bases or polynucleotides containing
PT them.
XX
XX Claim 6; SEQ ID NO 227; 423pp; French.
XX
XX The invention relates to the entire genome of cyanophaga S-2L, and to the
CC protein encoded by it. Genes isolated from the genome of S-2L are useful
CC for preparing enzymes for synthesis of D-bases (D = 2,6-diaminopurine),
CC particularly D, dTMP and dTTP, or polynucleotides containing these bases,
CC polymerases involved in metabolism of D-bases and deoxynucleotide
XX

```

CC analogs, for chemotherapy of AIDS. The genes, and encoded polypeptides,
CC can be used for detection and/or identification of S-2L, and for
CC identifying agents that modulate synthesis of D-bases or polynucleotides
CC containing them, and fusions of S-2L polypeptides with an antigen can be
CC used to raise specific antibodies, useful for detecting S-2L. This
CC sequence corresponds to one of the proteins encoded by the cyanophage S-
CC 2L genome.

Sequence 166 AA;

Query Match 59.4%; Score 41; DB 8; Length 166;
Best Local Similarity 53.8%; Pred. NO. 1.1e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EGGPGTTSNRLDA 13
 :||| ::||
Dd 36 OGGPGAGADPRADA 48

RESULT 8

ABO74099
ID ABO74099 standard; protein; 305 AA.

AC ABO74099:

29-JUL-2004 (first entry)

XX
DE Pseudomonas aeruginosa polypeptide #6274.
XX

XX Bacterial infection; *Pseudomonas aeruginosa* infection; antibacterial.

OS *Pseudomonas aeruginosa*.

US6551795-B1.

22-APR-2003

XX
PF
18-FEB-1999.

XX
18-FEB-1998. 0911S-0074798D

PR 27-JUL-1998; 98US-0094190P.
vv

PA (GENO-) GENOME THERAPEUTICS CORP.

PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;

DR WPI; 2003-615309/58.

DR N-PSDB: ABD07670;

Novel isolated nucleic acid encoding *Pseudomonas aeruginosa* polypeptide, useful as molecular targets for diagnostics, prophylaxis and treatment of pathological conditions resulting from bacterial infection

PS Disclosure: SEQ ID NO 22845; 455pp; English;

The invention relates to *Pseudomonas aeruginosa* polypeptides and the polynucleotides encoding them. The sequences are useful in diagnosis and therapy of pathological conditions, as molecular targets for diagnostics, prophylaxis and treatment of pathological conditions resulting from a bacterial infection, for evaluating a compound, such as a polypeptide, for the ability to bind a *P. aeruginosa* nucleic acid, as components of effective anti-bacterial targets, as targets for antibacterial drugs, including anti-*P. aeruginosa* drugs, as templates for recombinant production of *P. aeruginosa*-derived peptides or polypeptides, as target components for diagnosis and/or treatment of *P. aeruginosa*-caused infection, and in detection of *P. aeruginosa* sequences or other sequences of *Pseudomonas* species using biochip technology. Sequences AB067826-AB084396 represent *P. aeruginosa* polypeptides of the invention. Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format from USPTO at seqdata.uspto.gov/sequence.html

Sequence 305 AA:

Query Match 59.4%; Score 41; DB 7; Length 305;
Best Local Similarity 58.3%; Pred. No. 2.1e+02;
Matches 7; Conservative 3; Mismatches 2; Indels

Qy 1 EGGPGTTSNRLD 12
:|:|:|:|:|:|
Db 224 QGAPGTGASRLD 23

RESULT 9

ADN22235
ID ADN22235 standard; protein; 398 AA.

AC ADN22235:

XX
DT 02-DEC-20

XX
DE Postcard 1 vol. mont. 120 #498

XXII

KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.

OS Bacteria.

US2003233675-A1.

18-DEC-2003

XX
DE
20-FEB-2003. 2003US-00369493[illegible][illegible]

PA (HINK/) HINKLE G J.

PA (CHEN/) CHEN X.

XX

PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

WPI: 2004-061375/06

New recombinant DNA construct comprising a promoter positioned to provide for expression of a polynucleotide encoding a polypeptide from a microbial source, useful for producing plants with improved properties.

PS Claim 1: SEO ID NO 4888: 122pp: English.

The invention relates to a recombinant DNA construct comprising a promoter functional in a plant cell, where the promoter is positioned to provide for expression of a polynucleotide encoding a polypeptide from a microbial source. The invention also relates to a transformed plant comprising the recombinant DNA construct and a method of producing a transformed plant having an improved property. The plant is a crop plant such as maize or soybean. The method of producing a transformed plant having an improved property comprises transforming a plant with the recombinant DNA construct and growing the transformed plant, where the polynucleotide or polypeptide is useful for improving plant properties. The recombinant DNA construct is useful for producing plants with improved plant properties, e.g. improved cold, heat or drought tolerance, tolerance to herbicides, extreme osmotic conditions, pathogens or pests, increased resistance to plant disease, better growth rate by modification of the cell cycle pathway with plant growth regulators, increased rate of homologous recombination, modified seed oil or protein yield and/or content, improved yield by modification of carbohydrate, nitrogen or phosphorus use and/or uptake, by modification of photosynthesis or providing improved plant growth and development under at least one stress condition.


```
QY      2 GGGGTTNRLDA 13
DB      4312 GGGGTAEPERDA 4323
|||||
ADG22495
ID ADG22495 standard; protein; 92 AA.
AC ADG22495;
XX
XX DT 26-FEB-2004 (first entry)
XX
XX DE Cyanophage S-2L encoded protein #240.
XX
XX KW genome; cyanophage; 2; 6-diaminopurine; chemotherapy; AIDS.
XX
XX OS Cyanophage S-2L.
XX
XX PN FR2839079-A1.
XX
XX PD 31-OCT-2003.
XX
XX PF 30-APR-2002; 2002FR-00005424.
XX
XX PR 30-APR-2002; 2002FR-00005424.
XX
XX (INSP ) INST PASTEUR.
XX
XX PA (CNRS ) CNRS CENT NAT RECH SCI.
XX
XX PA (GENO-) GENOSCOPE CENT NAT SEQUENCAGE GRP INTERE.
XX
XX PI Marliere P, Kaminski PA, Gallisson F, Rouzon M, Pochet S;
XX
XX PI Weissenbach J, Saurin W, Robert C, Vico V;
XX
XX DR WPI; 2004-045746/05.
XX
XX DR N-PSDB; ADG22255.
XX
XX PT New genomic sequence for cyanophage S-2L, useful for identifying genes
XX
XX PT for synthesis of 2,6-diaminopurine bases or polynucleotides containing
XX
XX PS Claim 6; SEQ ID NO 241; 423pp; French.
XX
XX CC The invention relates to the entire genome of cyanophage S-2L, and to the
XX
XX CC protein encoded by it. Genes isolated from the genome of S-2L are useful
XX
XX CC for preparing enzymes for synthesis of D-bases (D = 2,6-diaminopurine),
XX
XX CC particularly D, dDMP and dDTP, or polynucleotides containing these bases,
XX
XX CC polymerases involved in metabolism of D-bases and deoxynucleotide
XX
XX CC analogs, for chemotherapy of AIDS. The genes, and encoded polypeptides,
XX
XX CC can be used for detection and/or identification of S-2L, and for
XX
XX CC identifying agents that modulate synthesis of D-bases or polynucleotides
XX
XX CC containing them, and fusions of S-2L polypeptides with an antigen can be
XX
XX CC used to raise specific antibodies, useful for detecting S-2L. This
XX
XX CC sequence corresponds to one of the proteins encoded by the cyanophage S-
XX
XX CC 2L genome.
XX
XX SQ Sequence 92 AA;
Query Match 58.0%; Score 40; DB 8; Length 92;
Best Local Similarity 77.8%; Pred. No. 90;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY      2 GGGGTTNRL 10
DB      43 GGGGTTSGR 51
|||||
RESULT 13
AAY74086
ID AAY74086 standard; protein; 224 AA.
XX
XX AC AAY74086;
```

```
XX
XX DT 14-MAR-2000 (first entry)
XX
XX DE Human prostate tumor EST fragment derived protein #273.
XX
XX KW Pancreas; tumor; EST; expressed sequence tag; human; cytostatic;
XX
XX OS Homo sapiens.
XX
XX PN DE19820190-A1.
XX
XX PD 04-NOV-1999.
XX
XX PF 28-APR-1998; 98DE-01020190.
XX
XX PR 28-APR-1998; 98DE-01020190.
XX
XX PA (META-) METAGEN GES GENOMFORSCHUNG MBH.
XX
XX PI Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl B;
XX
XX DR WPI; 1999-621386/54.
XX
XX DR N-PSDB; AAZ52948.
XX
XX PT New human nucleic acid sequences from pancreatic tumors, and related
XX
XX PT proteins.
XX
XX PS Claim 23; Page 422; 502pp; German.
XX
XX CC This invention describes novel polypeptides and their encoding nucleic
XX
XX CC acids derived from human pancreatic tumor tissue which have cytostatic
XX
XX CC activity. The sequences are also useful in producing pharmaceutical
XX
XX CC compositions for treatment of pancreatic tumors. AAY73814-Y74252
XX
XX CC represent protein fragments encoded by the human pancreatic tumor cDNA
XX
XX CC library derived expressed sequence tag (EST) sequences represented in
XX
XX CC AAZ52858-253014
XX
XX SQ Sequence 224 AA;
Query Match 58.0%; Score 40; DB 2; Length 224;
Best Local Similarity 77.8%; Pred. No. 2.3e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY      1 EGGPGTTSN 9
DB      91 EGGPGTSSN 99
|||||
RESULT 14
AAM85724
ID AAM85724 standard; protein; 77 AA.
XX
XX AC AAM85724;
XX
XX DT 07-NOV-2001 (first entry)
XX
XX DE Human immune/haematopoietic antigen SEQ ID NO:13317.
XX
XX KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX
XX KW cytostatic; gene therapy; vaccine; metastasis.
XX
XX OS Homo sapiens.
XX
XX PN WO200157182-A2.
XX
XX PD 09-AUG-2001.
XX
XX PF 17-JAN-2001; 2001WO-US001354.
XX
XX PR 31-JAN-2000; 2000US-0179065P.
XX
XX PR 04-FEB-2000; 2000US-0180628P.
XX
XX PR 24-FEB-2000; 2000US-0184664P.
```

PR 02-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198123P.
 PR 19-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 07-JUL-2000; 2000US-0216880P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 11-JUL-2000; 2000US-0217496P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225477P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226686P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 01-SEP-2000; 2000US-0229349P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 08-SEP-2000; 2000US-0232081P.
 PR 12-SEP-2000; 2000US-0231968P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 26-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235836P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.

PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239335P.
 PR 13-OCT-2000; 2000US-0239337P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 PR (HUMA-) HUMAN GENOME SCI INC.
 PR Rosen CA, Barash SC, Ruben SM;
 PR WPI; 2001-483426/52.
 PR N-PSDB; AAK58505.
 PR Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PR useful for preventing, diagnosing and/or treating cancers and metastasis.
 PR Claim 11; SEQ ID NO 13317; 3071pp + Sequence Listing; English.
 PR XX
 PR XX
 PR PI
 PR XX
 PR XX
 PR DR
 PR XX
 PR XX
 PR PT
 PR XX
 PR PS
 PR XX

CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
 CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patients own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting the
 CC nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/haematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/haematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
 CC represent sequences used in the exemplification of the present invention
 XX
 SQ Sequence 77 AA;

Query Match 56.5%; Score 39; DB 4; Length 77;
 Best Local Similarity 63.8%; Pred. No. 1.1e+02;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 EGGPGTTSNRL 11
 ||||| |:
 Db 47 EGGPGRASRV 57

RESULT 15
 ABU41557
 ID ABU41557 standard; protein; 130 AA.

AC ABU41557;

XX 19-JUN-2003 (first entry)

XX Protein encoded by Prokaryotic essential gene #27084.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX Pseudomonas syringae.

XX WO200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US0009107.

XX 21-MAR-2001; 2001US-00815242.

XX 06-SEP-2001; 2001US-00948993.

XX 25-OCT-2001; 2001US-0342923P.

XX 08-FEB-2002; 2002US-00072851.

XX 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

XX Wall D, Trawick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.

XX N-PSDB; ACA45427.

XX New antisense nucleic acids, useful for identifying proteins or screening

XX for homologous nucleic acids required for cellular proliferation to

XX isolate candidate molecules for rational drug discovery programs.

XX Claim 25; SEQ ID NO 69481; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of

XX the 6213 antisense sequences given in the specification where expression

CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC the gene product; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 130 AA;

Query Match 56.5%; Score 39; DB 6; Length 130;
 Best Local Similarity 61.5%; Pred. No. 1.9e+02;
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 EGGPGTTSNRLDA 13

Db 11 EGGPGSTPRRLGA 23

Search completed: April 28, 2005, 18:17:31
 Job time : 100.906 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2005, 18:06:50 ; Search time 26.4062 Seconds
(without alignments)
36.750 Million cell updates/sec

Title: US-10-088-639A-2_COPY_226_238
Perfect score: 69
Sequence: 1 EGGPGTTSNRDLA 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/1aa/5A-COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B-COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A-COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B-COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTUS-COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	59.4	305	4	US-09-252-991A-22845
2	39	56.5	219	4	US-09-893-737-130
3	39	56.5	351	4	US-09-252-991A-17299
4	39	56.5	461	4	US-09-134-000C-5551
5	38	55.1	440	4	US-09-252-991A-27594
6	38	55.1	610	4	US-09-252-991A-20299
7	38	55.1	818	4	US-09-252-991A-16691
8	37	53.6	198	3	US-09-188-930-138
9	37	53.6	285	4	US-09-312-283C-138
10	37	53.6	285	4	US-09-312-283C-382
11	37	53.6	285	4	US-09-552-204A-2
12	37	53.6	294	3	US-09-188-930-294
13	37	53.6	294	4	US-09-312-283C-294
14	37	53.6	441	4	US-09-248-796A-14545
15	37	53.6	655	1	US-07-736-178C-2
16	37	53.6	1003	4	US-09-252-991A-24069
17	37	53.6	1287	4	US-09-252-991A-29606
18	36	52.2	159	3	US-08-415-655-7
19	36	52.2	215	4	US-09-902-540-10404
20	36	52.2	229	4	US-09-134-000C-4443
21	36	52.2	266	3	US-09-199-637A-157
22	36	52.2	275	4	US-09-252-991A-29227
23	36	52.2	300	3	US-09-064-703-10
24	36	52.2	305	3	US-09-064-703-11
25	36	52.2	325	4	US-09-252-991A-23098
26	36	52.2	417	3	US-08-640-906-4
27	36	52.2	417	3	US-08-640-906-18

28	36	52.2	417	3	US-09-395-936-4	Sequence 4, Appli
29	36	52.2	417	3	US-09-395-936-18	Sequence 18, Appli
30	36	52.2	420	4	US-09-252-991A-26653	Sequence 26653, A
31	36	52.2	424	4	US-09-252-991A-30209	Sequence 30209, A
32	36	52.2	426	4	US-09-489-039A-13845	Sequence 13845, A
33	36	52.2	441	4	US-09-540-236-3622	Sequence 3622, Ap
34	36	52.2	502	4	US-09-328-352-5891	Sequence 5891, Ap
35	36	52.2	568	4	US-09-489-039A-13755	Sequence 13755, A
36	36	52.2	575	4	US-09-449-632-5	Sequence 5, Appli
37	36	52.2	643	2	US-08-551-356-4	Sequence 4, Appli
38	36	52.2	643	5	PCT-US93-12687-4	Sequence 2, Appli
39	36	52.2	644	1	US-08-206-176-2	Sequence 2, Appli
40	36	52.2	644	4	US-09-919-039-121	Sequence 121, App
41	36	52.2	750	3	US-08-813-829-2	Sequence 2, Appli
42	36	52.2	750	3	US-08-813-829-2	Sequence 2, Appli
43	36	52.2	847	4	US-09-373-157-4	Sequence 4, Appli
44	36	52.2	875	4	US-09-949-016-8582	Sequence 8582, Ap
45	36	52.2	1203	4	US-09-769-699-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-252-991A-22845
; Sequence 22845, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22845
; LENGTH: 305
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22845

Query Match 59.4%; Score 41; DB 4; Length 305;
Best Local Similarity 58.3%; Pred. No. 47;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EGGPGTTSNRDL 12
Db 224 QGAPGTGASRLD 235
:||||:||||

RESULT 2
US-09-893-737-130
; Sequence 130, Application US/09893737
; Patent No. 6822082
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Presnell, Scott R.
; TITLE OF INVENTION: MAMMALIAN SECRETED PROTEINS
; FILE REFERENCE: 00-41
; CURRENT APPLICATION NUMBER: US/09/893,737
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US 60/215,446
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 329
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 130
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Homo sapiens

US-09-893-737-130

Query Match 56.5%; Score 39; DB 4; Length 219;
Best Local Similarity 46.2%; Pred. No. 71;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 EGGPGTTSNRDLA 13
:|||||:|
DB 33 QGGPGSTNSKQRA 45

RESULT 3

US-09-252-991A-17299
; Sequence 17299, Application US/09252991A
; Patent No. 6551795

; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 17299

; LENGTH: 351

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-17299

Query Match 56.5%; Score 39; DB 4; Length 351;
Best Local Similarity 77.8%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GGGPGTTSNR 10
|||||
DB 14 GGGPGTLNR 22

RESULT 4

US-09-134-000C-5551
; Sequence 5551, Application US/09134000C
; Patent No. 6617156

; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 032796-032

; CURRENT APPLICATION NUMBER: US/09/134,000C

; CURRENT FILING DATE: 1998-08-13

; PRIOR APPLICATION NUMBER: US 60/055,778

; PRIOR FILING DATE: 1997-08-15

; NUMBER OF SEQ ID NOS: 6812

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 5551

; LENGTH: 461

; TYPE: PRT

; ORGANISM: Enterococcus faecalis

US-09-134-000C-5551

Query Match 56.5%; Score 39; DB 4; Length 461;
Best Local Similarity 80.0%; Pred. No. 1.6e+02;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GGGPGTTSNR 11
|||||
DB 159 GRPGTFSNRL 168

RESULT 5

US-09-252-991A-27594

; Sequence 27594, Application US/09252991A
; Patent No. 6551795

; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 27594

; LENGTH: 440

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-27594

Query Match 55.1%; Score 38; DB 4; Length 440;
Best Local Similarity 70.0%; Pred. No. 2.2e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 GGGPGTTSNRL 11
|||||
DB 263 GGGPGPARNRL 272

RESULT 6

US-09-252-991A-20299

; Sequence 20299, Application US/09252991A
; Patent No. 6551795

; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 20299

; LENGTH: 610

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-20299

Query Match 55.1%; Score 38; DB 4; Length 610;
Best Local Similarity 58.3%; Pred. No. 3.2e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EGGPGTTSNRLD 12
|||||:|
DB 394 EGGPDTPHRL 405

RESULT 7

US-09-252-991A-16691

; Sequence 16691, Application US/09252991A
; Patent No. 6551795

; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 16691

; LENGTH: 405

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-16691

```
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 16691
; LENGTH: 818
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-16691

Query Match          55.1%; Score 38; DB 4; Length 818;
Best Local Similarity 61.5%; Pred. No. 4.4e+02;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY      1 EGGPGTTSNRLDA 13
      ||| | : |||
Db      718 EGGAGEQNRLDA 730

RESULT 8
US-09-188-930-138
; Sequence 138, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods For Their Use
; FILE REFERENCE: 11000.1011c1
; CURRENT APPLICATION NUMBER: US/09/188,930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 138
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Rat
US-09-188-930-138

Query Match          53.6%; Score 37; DB 3; Length 198;
Best Local Similarity 70.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 EGGPGTTSNR 10
      ||| | |||
Db      90 EGGPGRTGNR 99

RESULT 9
US-09-312-283C-138
; Sequence 138, Application US/09312283C
; Patent No. 6573095
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated from Skin Cells
; TITLE OF INVENTION: and Methods for Their Use
; FILE REFERENCE: 11000.1011c2
; CURRENT APPLICATION NUMBER: US/09/312,283C
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 425
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 138
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Mouse
US-09-312-283C-138

Query Match          53.6%; Score 37; DB 4; Length 198;
Best Local Similarity 70.0%; Pred. No. 2e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 EGGPGTTSNR 10
      ||| | |||
Db      81 EGGPGRTGNR 90

RESULT 10
US-09-312-283C-382
; Sequence 382, Application US/09312283C
; Patent No. 6573095
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated from Skin Cells
; TITLE OF INVENTION: and Methods for Their Use
; FILE REFERENCE: 11000.1011c2
; CURRENT APPLICATION NUMBER: US/09/312,283C
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 425
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 382
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Mouse
US-09-312-283C-382

Query Match          53.6%; Score 37; DB 4; Length 285;
Best Local Similarity 70.0%; Pred. No. 2e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 EGGPGTTSNR 10
      ||| | |||
Db      81 EGGPGRTGNR 90

RESULT 11
US-09-552-204A-2
; Sequence 2, Application US/09552204A
; Patent No. 6620909
; GENERAL INFORMATION:
; APPLICANT: Piddington, Christopher S.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOG ZACRP2
; FILE REFERENCE: 99-08
; CURRENT APPLICATION NUMBER: US/09/552,204A
; CURRENT FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: 60/130,207
; PRIOR FILING DATE: 1999-04-20
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-552-204A-2

Query Match          53.6%; Score 37; DB 4; Length 285;
Best Local Similarity 70.0%; Pred. No. 2e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 EGGPGTTSNR 10
      ||| | |||
Db      81 EGGPGRTGNR 90
```

RESULT 12
US-09-188-930-294
; Sequence 294, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods For Their Use
; FILE REFERENCE: 11000.1011c1
; CURRENT APPLICATION NUMBER: US/09/188,930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 294
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Rat
US-09-188-930-294

Query Match 53.6%; Score 37; DB 3; Length 294;
Best Local Similarity 70.0%; Pred. No. 2.1e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EGGPGTTSNR 10
||| ||| |||
DB 90 EGGPGRTGNR 99

RESULT 13
US-09-312-283C-294
; Sequence 294, Application US/09312283C
; Patent No. 6573095
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated from Skin Cells
; TITLE OF INVENTION: and Methods for Their Use
; FILE REFERENCE: 11000.1011c2
; CURRENT APPLICATION NUMBER: US/09/312,283C
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 425
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 294
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Mouse
US-09-312-283C-294

Query Match 53.6%; Score 37; DB 4; Length 294;
Best Local Similarity 70.0%; Pred. No. 2.1e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EGGPGTTSNR 10
||| ||| |||
DB 90 EGGPGRTGNR 99

RESULT 14
US-09-248-796A-14545
; Sequence 14545, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN

; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 14545
; LENGTH: 441
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-14545

Query Match 53.6%; Score 37; DB 4; Length 441;
Best Local Similarity 75.0%; Pred. No. 3.3e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 GGGGTTSN 9
||| ||| |||
DB 404 GGGGTGN 411

RESULT 15
US-07-736-178C-2
; Sequence 2, Application US/07736178C
; Patent No. 5514544
; GENERAL INFORMATION:
; APPLICANT: Rao, Ramachandra N
; APPLICANT: Turner, Jan R
; TITLE OF INVENTION: ACTIVATOR GENE FOR MACROLIDE
; TITLE OF INVENTION: BIOSYNTHESIS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly And Company
; STREET: Lilly corporate Center
; CITY: Indianapolis
; STATE: IN
; COUNTRY: USA
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/736,178C
; FILING DATE: 19910726
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Conrad, Robert A
; REGISTRATION NUMBER: 32089
; REFERENCE/DOCKET NUMBER: X8144
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3334
; TELEFAX: 317-276-1294
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 655 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-736-178C-2

Query Match 53.6%; Score 37; DB 1; Length 655;
Best Local Similarity 72.7%; Pred. No. 5.1e+02;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 GGGGTTSNR 12
||| ||| |||
DB 436 GGGGTTSNR 446

Search completed: April 28, 2005, 18:28:10
Job time : 28.4062 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2005, 18:24:46 ; Search time 80.6406 Seconds
(without alignments)
53.700 Million cell updates/sec

Title: US-10-088-639A-2_COPY_226_238

Perfect score: 69
Sequence: 1 EGGPTTSNRDLA 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1426032 seqs, 333106140 residues

Total number of hits satisfying chosen parameters: 1426032

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Published Applications, AA.*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	63.8	3149	16 US-10-408-765A-2326	Sequence 2326, Ap
2	42	60.9	566	15 US-10-424-599-166000	Sequence 166000,
3	41	59.4	398	15 US-10-369-493-4888	Sequence 4888, Ap
4	41	59.4	400	15 US-10-369-493-7646	Sequence 7646, Ap
5	41	59.4	19608	15 US-10-084-846A-8	Sequence 8, Appli
6	39	56.5	69	15 US-10-424-599-163360	Sequence 163360,
7	39	56.5	130	15 US-10-424-599-206460	Sequence 206460,
8	39	56.5	85	15 US-10-282-122A-69481	Sequence 69481, A
9	39	56.5	188	16 US-10-767-701-39568	Sequence 39568, A
10	39	56.5	203	16 US-10-437-963-130792	Sequence 130792,
11	39	56.5	219	9 US-09-893-737-130	Sequence 130, App
12	39	56.5	385	14 US-10-156-761-12542	Sequence 12542, A
13	39	56.5	454	9 US-09-815-242-10700	Sequence 10700, A

ALIGNMENTS

RESULT 1

US-10-408-765A-2326

Sequence 2326, Application US/10408765A
Publication No. US20040101874A1
GENERAL INFORMATION:
APPLICANT: Ghosh, Soumitra S.
APPLICANT: Fahy, Eoin D.
APPLICANT: Zhang, Bing
APPLICANT: Gibson, Bradford W.
APPLICANT: Taylor, Steven W.
APPLICANT: Glenn, Gary M.
APPLICANT: Warnock, Dale E.
TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
FILE REFERENCE: 660088.465
CURRENT APPLICATION NUMBER: US/10/408, 765A
CURRENT FILING DATE: 2003-04-04
NUMBER OF SEQ ID NOS: 3077
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 2326
LENGTH: 3149
TYPE: PRT
ORGANISM: Homo sapiens
US-10-408-765A-2326

Query Match 63.8%; Score 44; DB 16; Length 3149;

Best Local Similarity 80.0%; Pred. No. 8.4e+02; Mismatches 1; Indels 0; Gaps 0;

QY 3 GPGTTSNRDL 12

Db 72 GPGTTSNRDL 81

RESULT 2

US-10-424-599-166000

Sequence 57133, A
Sequence 45, Appl
Sequence 18, Appl
Sequence 797, App
Sequence 797, App
Sequence 141326,
Sequence 786, App
Sequence 786, App
Sequence 18, Appl
Sequence 248992,
Sequence 195589,
Sequence 3825, A
Sequence 6547, A
Sequence 1149, Ap
Sequence 159302,
Sequence 14175, A
Sequence 3298, Ap
Sequence 8193, Ap
Sequence 43771, A
Sequence 43208, A
Sequence 150859,
Sequence 122878,
Sequence 171832,
Sequence 2611, Ap
Sequence 67846, A
Sequence 138, App
Sequence 382, App
Sequence 5, Appli
Sequence 60, Appl
Sequence 88, Appl
Sequence 2, Appli
Sequence 294, App

15 US-10-282-122A-57133
15 US-10-182-243-45
15 US-10-415-011-18
9 US-09-925-302-797
10 US-09-925-302-797
16 US-10-437-963-141326
9 US-09-925-298-786
14 US-10-102-806-786
10 US-09-981-151A-18
15 US-10-424-599-248992
16 US-10-437-963-195589
16 US-10-767-701-3825
16 US-10-425-114-65647
15 US-10-369-493-1149
15 US-10-424-599-159302
14 US-10-156-761-14175
15 US-10-320-797-3298
15 US-10-369-493-8193
9 US-09-864-761-47371
9 US-09-864-761-43208
15 US-10-424-599-150859
16 US-10-437-963-122878
16 US-10-437-963-171832
146 US-10-276-774-2611
192 US-10-282-122A-67846
198 US-09-866-050A-138
10 US-09-866-050A-382
14 US-10-234-000-5
285 US-10-411-120-60
285 US-10-411-120-88
285 US-10-621-787-2
294 US-09-866-050A-294

454 56.5 39 56.5
600 56.5 39 56.5
600 56.5 39 56.5
609 56.5 39 56.5
609 56.5 39 56.5
97 55.1 38 55.1
102 55.1 38 55.1
102 55.1 38 55.1
134 55.1 38 55.1
134 55.1 38 55.1
151 55.1 38 55.1
152 55.1 38 55.1
213 55.1 38 55.1
248 55.1 38 55.1
293 55.1 38 55.1
358 55.1 38 55.1
409 55.1 38 55.1
803 55.1 38 55.1
1090 55.1 38 55.1
56 53.6 37 53.6
58 53.6 37 53.6
96 53.6 37 53.6
106 53.6 37 53.6
120 53.6 37 53.6
146 53.6 37 53.6
192 53.6 37 53.6
198 53.6 37 53.6
285 53.6 40 53.6
285 53.6 41 53.6
285 53.6 42 53.6
285 53.6 43 53.6
285 53.6 44 53.6
294 53.6 45 53.6

```

; Sequence 166000, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovacic David K
; APPLICANT: Zhou Yihwei
; APPLICANT: Cao Yungwei
; TITLE OF INVENTION: Soy Nucleic Acid Mo
; TITLE OF INVENTION: Plants and Uses Th
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,5
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285604
; SEQ ID NO 166000
; LENGTH: 566
; TYPE: PRNT
; ORGANISM: Glycine max
; FEATURE:
; US-10-424-599-166000
; OTHER INFORMATION: Clone ID: PAT_MRT38

```

Query Match	60.9%	Score 42;	DB 15;	Length 566;
Best Local Similarity	72.7%	Pred. NO. 3e+02;		
Matches	8;	Conservative	0;	Mismatches 3;
				Indels 0;
				Gaps 0;

Qy 2 GGPGTTSNRLD 12
Db 549 GGEGTTRNHLD 559

RESULT 3
US-10-369-493-4888
; Sequence 4888, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION.

; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.

```

/ APPLICANT: Chen, Xianfeng
/
/ TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
/
/ TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
/
/ FILE REFERENCE: 38-10(52052)B
/
/ CURRENT APPLICATION NUMBER: US/10/369,493
/
/ CURRENT FILING DATE: 2003-02-28
/
/ PRIOR APPLICATION NUMBER: US 60/360,039
/
/ PRIOR FILING DATE: 2002-02-21
/
/ NUMBER OF SEQ ID NOS: 47374
/
/ SEQ ID NO 4888
/
/ LENGTH: 398
/

```

; TYPE: PRI
 ; ORGANISM: Burkholderia fungorum
 US-10-369-493-4888

Query Match	59.4%	Score 41;	DB 15;
Best Local Similarity	72.7%	Pred. No. 3.1e+02;	
Matches	8;	Mismatches	2;
Conservative		Indels	0;

Qy 1 EGGPGTTSNRL 11
: | | | | | | |
Db 77 OGVPGTTVNRL 87.

RESULT 4
US-10-369-493-7646
; Sequence 7646, Application US/10369493
; Publication No. US20030233675A1

; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng

; GENERAL INFORMATION:

; Sequence 166000, Application US/10424599
 ; Publication No. US20040031072A1
 ; GENERAL INFORMATION:
 ;
 ; APPLICANT: La Rosa Thomas J
 ; APPLICANT: Kovalic David K
 ; APPLICANT: Zhou Yihua
 ; APPLICANT: Cao Yongwei
 ;
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21 (53223) B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 166000

```

; LENGTH: 566
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:

```

OTHER INFORMATION: Clone ID: PAT_MRT3847_120914C.1.pcp
US-10-424-599-166000

Query Match	60.9%	Score 42;	DB 15;	Length 566;
Best Local Similarity	72.7%	Pred. NO. 3e+02;		
Matches	8;	Conservative	0;	Mismatches 3;
				Indels 0;
				Gaps 0;

Qy 2 GPGTTSNRLD 12
Db 549 GGEGTTRNHL 559

RESULT 3
US-10-369-493-4888
; Sequence 4888, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION.

GENERAL INFORMATION:
 APPLICANT: Cao, Yongwei
 APPLICANT: Hinkle, Gregory J.
 APPLICANT: Slater, Steven C.
 APPLICANT: Goldman, Barry S.

```

/ APPLICANT: Chen, Xianfeng
/
/ TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
/
/ TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
/
/ FILE REFERENCE: 38-10(52052)B
/
/ CURRENT APPLICATION NUMBER: US/10/369,493
/
/ CURRENT FILING DATE: 2003-02-28
/
/ PRIOR APPLICATION NUMBER: US 60/360,039
/
/ PRIOR FILING DATE: 2002-02-21
/
/ NUMBER OF SEQ ID NOS: 47374
/
/ SEQ ID NO 4888
/
/ LENGTH: 398
/

```

; TYPE: PRI
 ; ORGANISM: Burkholderia fungorum
 US-10-369-493-4888

Query Match	59.4%	Score 41;	DB 15;
Best Local Similarity	72.7%	Pred. No. 3.1e+02;	
Matches 8;	Conservative	1;	Mismatches 2;
			Indels 0;
			Gaps 0;

Qy 1 EGGPGTTSNRL 11
: | | | | | | |
Db 77 OGVPGTTVNRL 87.

RESULT 4
US-10-369-493-7646
; Sequence 7646, Application US/10369493
; Publication No. US20030233675A1

; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng

; GENERAL INFORMATION:

```

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
;
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
;
; FILE REFERENCE: 38-10(52052)B
;
; CURRENT APPLICATION NUMBER: US/10/369,493
;
; CURRENT FILING DATE: 2003-02-28
;
; PRIOR APPLICATION NUMBER: US 60/360,039
;
; PRIOR FILING DATE: 2002-02-21
;
; NUMBER OF SEQ ID NOS: 47374
;
; SEQ ID NO 7646
;
; LENGTH: 400
;
; TYPE: PRT
;
; ORGANISM: Burkholderia cepacia
;
; US-10-369-493-7646

```

Query Match 59.4%; Score 41; DB 15; Length 400;
Best Local Similarity 72.7%; Pred. NO. 3.1e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EGGPGTTSNRL 11
: | | | | |
Db 79 OGVPGTTVNRL 89

```

RESULT 5
US-10-084-846A-8
; Sequence 8, Application US/10084846A
; Publication No. US20040006026A1
; GENERAL INFORMATION:
; APPLICANT: WEITNAUER, GABRIELE
; APPLICANT: MUHLNBERG, AGNES
; APPLICANT: TREFFNER, AXEL
; APPLICANT: BECHTHOLD, ANDREAS
; TITLE OF INVENTION: AVILAMYCIN DERIVATIVES
; FILE REFERENCE: 1974-005
; CURRENT APPLICATION NUMBER: US/10/084,846A
; CURRENT FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: PCT/EP01/09815
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: DE 101 09 166.4
; PRIOR FILING DATE: 2001-02-25
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 8
; LENGTH: 19608
; TYPE: PRT
; ORGANISM: Streptomyces viidochromogenes
; FEATURE:
; OTHER INFORMATION: Protein 3: amino acid
; OTHER INFORMATION: Start codon: atc, Star
US-10-084-846A-8

```

Query Match 59.4%; Score 41; DB 15; Length 19608;
Best Local Similarity 66.7%; Pred. No. 1.6e+04;
Matches 8: Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 GGPGTTSNRLDA 13
|||
Db 4261 GGPGTAEPRRDA 4272

```

RESULT 6
US-10-424-599-163360
; Sequence 163360, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Mo
; TITLE OF INVENTION: Plants and Uses Th
; FILE REFERENCE: 38-21(53223) B
; CURRENT APPLICATION NUMBER: US/10/424,5

```

; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 163360
; LENGTH: 69
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_118532C.1.pap
US-10-424-599-163360

Query Match 56.5%; Score 39; DB 15; Length 69;
Best Local Similarity 63.8%; Pred. No. 1.1e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EGGPGTTSNRL 11
||||| : |||
Db 42 EGGPQTSLNRM 52

RESULT 7

US-10-424-599-206460
; Sequence 206460, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 206460
; LENGTH: 85
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_28460C.1.pap
US-10-424-599-206460

Query Match 56.5%; Score 39; DB 15; Length 85;
Best Local Similarity 87.5%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GGGPTTSN 9
||||| : |||
Db 52 GGGPTTIN 59

RESULT 8

US-10-282-122A-69481
; Sequence 69481, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 69481
; LENGTH: 130
; TYPE: PRT
; ORGANISM: Pseudomonas syringae
US-10-282-122A-69481

Query Match 56.5%; Score 39; DB 15; Length 130;
Best Local Similarity 61.5%; Pred. No. 2e+02;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 EGGPGTTSNRLDA 13
||||| : |||
Db 11 EGTGSGTPRLCA 23

RESULT 9

US-10-767-701-39568
; Sequence 39568, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 39568
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C20858_1.pap
US-10-767-701-39568

Query Match 56.5%; Score 39; DB 16; Length 188;
Best Local Similarity 63.6%; Pred. No. 2.9e+02;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 GPGTTSNRLDA 13
||||| : |||
Db 109 GPSTTDEKIDA 119

RESULT 10

US-10-437-963-190792
; Sequence 190792, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 190792
; LENGTH: 203
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_87172C.1.pep
US-10-437-963-190792

Query Match 56.5%; Score 39; DB 16; Length 203;
Best Local Similarity 70.0%; Pred. No. 3.2e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 GPGTTSNRD 12
Db 71 GPGTSSHRSD 80

RESULT 11
US-09-893-737-130
; Sequence 130, Application US/09893737
; Patent No. US20020110855A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Presnell, Scott R.
; TITLE OF INVENTION: MAMMALIAN SECRETED PROTEINS
; FILE REFERENCE: 00-41
; CURRENT APPLICATION NUMBER: US/09/893,737
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US 60/215,446
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 329
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 130
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-893-737-130

Query Match 56.5%; Score 39; DB 9; Length 219;
Best Local Similarity 46.2%; Pred. No. 3.4e+02;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 EGGPGTTSNRD 13
Db 33 QGGPGSTNSKQRA 45

RESULT 12
US-10-156-761-12542
; Sequence 12542, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA

; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 12542
; LENGTH: 385
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-12542

Query Match 56.5%; Score 39; DB 14; Length 385;
Best Local Similarity 77.8%; Pred. No. 6.1e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 GPGGTTSNR 10
Db 364 GPGGTSSR 372

RESULT 13
US-09-815-242-10700
; Sequence 10700, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011a
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10700
; LENGTH: 454
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-815-242-10700

Query Match 56.5%; Score 39; DB 9; Length 454;
Best Local Similarity 80.0%; Pred. No. 7.2e+02;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GPGGTTSNR 11
Db 152 GPGGTSSNR 161

RESULT 14

US-10-282-122A-57133
; Sequence 57133, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034A

; CURRENT APPLICATION NUMBER: US/10/282,122A

; CURRENT FILING DATE: 2003-02-20

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/230,335

; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: 60/230,347

; PRIOR FILING DATE: 2000-09-09

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/267,636

; PRIOR FILING DATE: 2001-02-09

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 78614

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 57133

; LENGTH: 454

; TYPE: PRT

; ORGANISM: Enterococcus faecalis

US-10-282-122A-57133

Query Match 56.5%; Score 39; DB 15; Length 454;

Best Local Similarity 80.0%; Pred. No. 7.2e+02;

Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GGGTTSNRL 11

Db 152 GGGTTSNRL 161

RESULT 15

US-10-182-243-45

; Sequence 45, Application US/10182243

; Publication No. US20040048310A1

; GENERAL INFORMATION:

; APPLICANT: PLOWMAN, GREGORY D.

; APPLICANT: WHYTE, DAVID

; APPLICANT: MANNING, GERARD

; APPLICANT: SUDARSANAM, SUCHA

; APPLICANT: MARTINEZ, RICARDO

; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES AND PROTEIN KINASE-LIKE

; TITLE OF INVENTION: ENZYMES

; FILE REFERENCE: 038602/1366

; CURRENT APPLICATION NUMBER: US/10/182,243

; CURRENT FILING DATE: 2003-07-07

; PRIOR APPLICATION NUMBER: PCT/US01/02337

; PRIOR FILING DATE: 2001-01-25

; NUMBER OF SEQ ID NOS: 84

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 45

; LENGTH: 600

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-182-243-45

Query Match 56.5%; Score 39; DB 15; Length 600;

Best Local Similarity 87.5%; Pred. No. 9.5e+02;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EGGPGTTS 8

Db 187 EGGPGATS 194

Search completed: April 28, 2005, 19:02:17

Job time : 81.6406 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2005, 18:05:55 ; Search time 17.875 Seconds
(without alignments)
69.976 Million cell updates/sec

Title: US-10-088-639A-2_COPY_226_238
Perfect score: 69
Sequence: 1 EGGPGTTSNRDLA 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79: *
1: Pir1: *
2: Pir2: *
3: Pir3: *
4: Pir4: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	43	62.3	149	2 C87666	ribonuclease HI [i
2	40	58.0	249	2 AG0618	probable DNA repli
3	40	58.0	249	2 T03011	dnaC protein homol
4	40	58.0	344	2 S57901	estradiol 17beta-d
5	40	58.0	351	2 T36010	hypothetical prote
6	39	56.5	1073	2 T33764	hypothetical prote
7	39	56.5	1112	2 H95964	probable outer mem
8	39	56.5	3938	2 T42761	Bassoon protein -
9	38	55.1	293	2 E69174	succinate-CoA liga
10	38	55.1	293	2 F83618	conserved hypothet
11	38	55.1	312	2 AG2884	hypothetical prote
12	38	55.1	312	2 D97660	hypothetical prote
13	38	55.1	352	2 F70134	flagellar motor sw
14	38	55.1	667	2 F70882	probable membrane
15	38	55.1	1872	2 T24683	hypothetical prote
16	37	53.6	139	2 E75346	organic hydroperox
17	37	53.6	194	2 I57523	HSP90 - mouse (fra
18	37	53.6	222	2 A72643	hypothetical prote
19	37	53.6	229	2 T40148	hypothetical prote
20	37	53.6	290	2 T03552	maltose transport
21	37	53.6	353	2 T09665	peroxidase (EC 1.1
22	37	53.6	385	2 H86731	conserved hypothet
23	37	53.6	411	2 S45318	keratin 12 - rabbi
24	37	53.6	443	2 T21499	hypothetical prote
25	37	53.6	604	2 S25203	smr protein - Str
26	37	53.6	616	2 A72627	probable 2-oxoacid
27	36	52.2	143	2 A49438	p53 tumor suppress
28	36	52.2	148	2 A83418	ribonuclease H PA1
29	36	52.2	159	2 I49023	tumor suppressor p

30	36	52.2	240	2 A82783	hypothetical prote
31	36	52.2	313	2 T33010	hypothetical prote
32	36	52.2	330	2 E81717	Holliday junction
33	36	52.2	334	2 C71565	DNA-binding protei
34	36	52.2	341	2 T08459	hypothetical prote
35	36	52.2	341	2 T48847	synthetase synt4 [im
36	36	52.2	351	2 JC4779	peroxidase (EC 1.1
37	36	52.2	379	2 T40384	probable methionin
38	36	52.2	400	2 E96029	probable beta-keto
39	36	52.2	417	2 A56171	carboxypeptidase A
40	36	52.2	471	2 B86170	ADK1 [imported] -
41	36	52.2	549	2 E83085	conserved hypothet
42	36	52.2	582	2 T23624	hypothetical prote
43	36	52.2	617	2 T23623	hypothetical prote
44	36	52.2	644	1 FGHUA	fibrinogen alpha c
45	36	52.2	668	2 E84253	chemotaxis protein

ALIGNMENTS

RESULT 1

C87666
ribonuclease HI [imported] - Caulobacter crescentus
C/Species: Caulobacter crescentus
C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C/Accession: C87666
R/Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A/Title: Complete Genome Sequence of Caulobacter crescentus.
A/Reference number: A87249; MUID:21173698; PMID:11259647
A/Accession: C87666
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-149 <STO>
A/Cross-references: UNIPROT:Q9A341; GB:AE005673; NID:gl3425071; PIDN:AAK25327.1; GSPDB:G
C/Genetics:
A/Gene: CC3365
C/Superfamily: ribonuclease H

RESULT 2

AG0618
probable DNA replication protein STY1023 [imported] - Salmonella enterica subsp. enterica
C/Species: Salmonella enterica subsp. enterica serovar Typhi
A/Note: this species has also been called Salmonella typhi
C/Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C/Accession: AG0618
R/Farkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A/Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A/Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A/Reference number: AB0502; MUID:21534947; PMID:11677608
A/Accession: AG0618
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-249 <PAR>
A/Cross-references: GB:AL513382; PIDN:CAD05417.1; PID:gl6502178; GSPDB:GN00176
C/Genetics:
A/Gene: STY1023
C/Superfamily: DNA replication protein dnaC

Query Match 58.0%; Score 40; DB 2; Length 249;
 Best Local Similarity 66.7%; Pred. No. 21;
 Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 2 GPGGTTSNRLDA 13
 ||||| |
 Db 109 GPGGTGNHLAA 120

RESULT 3
 T03011
 dnaC protein homolog - Salmonella typhimurium
 C:Species: Salmonella typhimurium
 C>Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 09-Jul-2004
 C:Accession: T03011
 R:Figueroa-Bossi, N.; Bossi, L.
 submitted to the EMBL Data Library, June 1998
 A:Reference number: Z14818
 A:Accession: T03011
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-249 <FIG>
 A:Cross-references: UNIPROT:O84893; EMBL:AF001386; NID:G3294471; PIDN:AAC26072.1; PID:93
 C:Superfamily: DNA replication protein dnaC

Query Match 58.0%; Score 40; DB 2; Length 249;
 Best Local Similarity 66.7%; Pred. No. 21;
 Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 2 GPGGTTSNRLDA 13
 ||||| |
 Db 109 GPGGTGNHLAA 120

RESULT 4
 S57901
 estradiol 17beta-dehydrogenase (EC 1.1.1.62) - rat
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
 C:Accession: S57901
 R:Ghersevich, S.; Nokelainen, P.; Poutanen, M.; Orava, M.; Autio-Harmainen, H.; Rajaniemi
 Endocrinology 135, 1477-1487, 1994
 A:Title: Rat 17-beta-hydroxysteroid dehydrogenase type 1: primary structure and regulati
 A:Reference number: S57901; MUID:95009707; PMID:7925110
 A:Accession: S57901
 A:Status: preliminary; nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-344 <GHE>
 A:Cross-references: UNIPROT:P51657; EMBL:X78811; NID:G561532; PIDN:CAA55389.1; PID:G5615
 C:Superfamily: 17beta-dehydrogenase; short-chain alcohol dehydrogenase homology
 C:Keywords: oxidoreductase
 F/4-189/Domain: short-chain alcohol dehydrogenase homology <SADH>

Query Match 58.0%; Score 40; DB 2; Length 344;
 Best Local Similarity 61.5%; Pred. No. 29;
 Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 EGGPGTTSNRLDA 13
 ||||| |
 Db 198 EGGPGGALERADA 210

RESULT 5
 T36010
 hypothetical protein SCC22.20 - Streptomyces coelicolor
 C:Species: Streptomyces coelicolor
 C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
 C:Accession: T36010
 R:Seeger, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrall, B.G.; Rajandream, M.A.
 submitted to the EMBL Data Library, July 1999
 A:Reference number: Z21574
 A:Accession: T36010

A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-351 <SEE>
 A:Cross-references: UNIPROT:Q9XAB8; EMBL:AL096839; PIDN:CAB50763.1; GSPDB:GN00070; SCOE
 A:Experimental source: strain A3(2)
 C:Genetics:
 A:Gene: SCOEDB:SCC22.20

Query Match 58.0%; Score 40; DB 2; Length 351;
 Best Local Similarity 63.6%; Pred. No. 30;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 GPGGTTSNRLD 12
 ||||| |
 Db 233 GPGGLTAVERMD 243

RESULT 6
 T33764
 hypothetical protein C02E11.1 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
 C:Accession: T33764
 R:Pauley, A.; Harper, M.
 submitted to the EMBL Data Library, October 1998
 A:Description: The sequence of C. elegans cosmid C02E11.
 A:Reference number: Z21403
 A:Accession: T33764
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-1073 <PAU>
 A:Cross-references: UNIPROT:Q8ITW0; UNIPROT:Q9UAN9; EMBL:AF101304; PIDN:AAC69200.1; GSPD
 A:Experimental source: strain Bristol N2; clone C02E11
 C:Genetics:
 A:Gene: CBSP:C02E11.1
 A:Map position: 5
 A:Introns: 37/1; 174/3; 248/3; 511/3; 702/2; 1043/3

Query Match 56.5%; Score 39; DB 2; Length 1073;
 Best Local Similarity 80.0%; Pred. No. 1.4e+02;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 GPGGTTSNRLD 12
 ||||| |
 Db 458 GPGTYSARLD 467

RESULT 7
 H95964
 Probable outer membrane secretion protein SMB21543 [imported] - Sinorhizobium meliloti (S
 C:Species: Sinorhizobium meliloti
 C>Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
 C:Accession: H95964
 R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan
 Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
 A:Title: The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-fixing endo
 A:Reference number: A95842; MUID:21396508; PMID:11481431
 A:Accession: H95964
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1112 <KUR>
 A:Cross-references: UNIPROT:Q92UV3; GB:AL591985; PIDN:CAC49384.1; PID:G15140870; GSPDB:G
 A:Experimental source: strain 1021, megaplasmid pSymb
 R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
 pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
 L.; Hyman, R.W.; Jones, T.
 Science 293, 668-672, 2001
 A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
 hebaud, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
 A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
 A:Reference number: A96039; MUID:21368234; PMID:11474104
 A:Contents: annotation
 C:Genetics:

A;Gene: SMB21543
A;Genome: plasmid

Query Match 56.5%; Score 39; DB 2; Length 1112;
Best Local Similarity 77.8%; Pred. No. 1.5e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 GGGPGTTSNR 10
| : : : : :
DB 913 GPGGTSNR 921

RESULT 8

T42761
Basoon protein - rat
N;Alternate names: brain-specific synapse-associated protein
C;Species: Rattus norvegicus (Norway rat)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C;Accession: T42761
R;Dieck, S.; Sammarti-Vila, L.; Langnaese, K.; Richter, K.; Kindler, S.; Soyke, A.; Wex, J. Cell Biol. 142, 499-509, 1998
A;Title: Basoon, a novel zinc-finger CAG/Glutamine-repeat protein selectively localized
A;Reference number: Z22249; MUID:198345363; PMID:9679147
A;Accession: T42761
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-3938 <DIE>
A;Cross-references: UNIPROT:O88778; EMBL:Y16563; NID:g3413503; PIDN:CAA76287.1; PID:g3413503
A;Experimental source: strain Sprague Dawley; brain
C;Function:
A;Description: may be involved in cytomatrix organization at the site of neurotransmitter release
A;Note: component of the presynaptic cytoskeleton
C;Keywords: coiled coil, zinc finger

Query Match 56.5%; Score 39; DB 2; Length 3938;
Best Local Similarity 70.0%; Pred. No. 5.7e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 GPGTTSNR 12
| : : : : :
DB 69 GPGSVRR 78

RESULT 9

E69174
succinate-CoA ligase (ADP-forming) (EC 6.2.1.5) alpha chain - Methanobacterium thermoautotrophicum
C;Species: Methanobacterium thermoautotrophicum
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 16-Aug-2004
C;Accession: E69174
R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.; Ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N. J. Bacteriol. 179, 7135-7155, 1997
A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: function
A;Reference number: A69000; MUID:98037514; PMID:9371463
A;Accession: E69174
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-293 <MTH>
A;Cross-references: UNIPROT:O26663; GB:AE000839; GB:AE000666; NID:g2621637; PIDN:AA88506
A;Experimental source: strain Delta H
C;Genetics:
A;Note: MTH563
C;Superfamily: Succinyl-CoA synthetase, alpha subunit
C;Keywords: acid-thiol ligase; ATP; coenzyme A; phosphohistidine; phosphoprotein; tricarboxylate; His (phosphohistidine intermediate) #status predicted
F;249/Active site: His (phosphohistidine intermediate) #status predicted

Query Match 55.1%; Score 38; DB 2; Length 293;
Best Local Similarity 53.8%; Pred. No. 55;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 EGGPGTTSNR 13
| : : : : :
DB 193 EGGAGEQNAR 205

DB 255 EGGTGTSKREA 267

RESULT 10

F83618
conserved hypothetical protein PA0209 [imported] - Pseudomonas aeruginosa (strain PA01)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: F83618
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Brj
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: F83618
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-293 <STO>
A;Cross-references: UNIPROT:O916S9; GB:AE004459; GB:AE004091; NID:g9946043; PIDN:AG0359
A;Experimental source: strain PA01
C;Genetics:
A;Gene: PA0209
C;Superfamily: citG protein

Query Match 55.1%; Score 38; DB 2; Length 293;
Best Local Similarity 61.5%; Pred. No. 55;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 EGGPGTTSNR 13
| : : : : :
DB 193 EGGAGEQNAR 205

RESULT 11

AG2884
hypothetical protein Atu2506 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C;Accession: AG2884
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.; erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClellan, Karp, P.; Romero, P.; Zhang, S. Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, E.; Star, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AG2884
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-312 <KUR>
A;Cross-references: UNIPROT:Q8UC11; GB:AE008688; PIDN:AAL43493.1; PID:g17741000; GSPDB:G
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: Atu2506
A;Map position: circular chromosome
C;Superfamily: inner membrane protein upga

Query Match 55.1%; Score 38; DB 2; Length 312;
Best Local Similarity 50.0%; Pred. No. 59;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 EGGPGTTSNR 12
| : : : : :
DB 252 GGGPGTASSETIN 263

RESULT 12

D97660
hypothetical protein AGR_C_4553 [imported] - Agrobacterium tumefaciens (strain C58, Cere
C;Species: Agrobacterium tumefaciens
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004

C;Accession: D97660
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorllo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tumefaciens* strain C58
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: D97660
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-312 <KUR>
A;Cross-references: UNIPROT:Q8UC11; GB:AE007869; PIDN:AAK88237.1; PID:g15157693; GSPDB:G0000021; CESP:T08A11.1
C;Genetics:
A;Gene: AGR_C_4553
A;Map position: circular chromosome
C;Superfamily: inner membrane protein ugpA

Query Match 55.1%; Score 38; DB 2; Length 312;
Best Local Similarity 50.0%; Pred. No. 59;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 EGGPGTTSNRLD 12
:|||||:
Db 252 QGGPGTASETIN 263

RESULT 13
F70134
flagellar motor switch protein (flm) homolog - Lyme disease spirochete
C;Species: *Borrelia burgdorferi* (Lyme disease spirochete)
C;Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 09-Jul-2004
C;Accession: F70134
R;Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt, J.; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B. Nature 390, 580-586, 1997
A;Authors: Smith, H.O.; Venter, J.C.
A;Title: Genomic sequence of a Lyme disease spirochete, *Borrelia burgdorferi*.
A;Reference number: A70100; MUID:98065943; PMID:9403685
A;Accession: F70134
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-352 <LE>
A;Cross-references: UNIPROT:Q57511; GB:AE001137; GB:AE000783; NID:g2688160; PIDN:AAK6667
A;Experimental source: strain B31
C;Superfamily: flagellar motor switch protein flm

Query Match 55.1%; Score 38; DB 2; Length 352;
Best Local Similarity 72.7%; Pred. No. 67;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 GPGTTSNRLD 13
:|||||:
Db 250 GVGTTSENLD 260

RESULT 14
F70682
probable membrane protein - *Mycobacterium tuberculosis* (strain H37RV)
C;Species: *Mycobacterium tuberculosis*
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C;Accession: F70682
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: F70682
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-667 <COL>
A;Cross-references: UNIPROT:P71749; GB:Z81368; GB:AL123456; NID:g3261656; PIDN:CAB03731.

A;Experimental source: strain H37RV
C;Genetics:
A;Accession: RV2395
C;Superfamily: *Pyrococcus horikoshii* hypothetical protein PH0361

Query Match 55.1%; Score 38; DB 2; Length 667;
Best Local Similarity 53.8%; Pred. No. 1.3e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 EGGPGTTSNRLD 13
:|||||:
Db 475 QGAPGATANALAA 487

RESULT 15
T24683
hypothetical protein T08A11.1 - *Caenorhabditis elegans*
C;Species: *Caenorhabditis elegans*
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T24683; T26904
R;Chui, C.; Sulston, J.
submitted to the EMBL Data Library, August 1995
A;Reference number: Z19922
A;Accession: T24683
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1872 <WIL>
A;Cross-references: UNIPROT:Q9XTE2; EMBL:Z50875; PIDN:CAA90776.1; GSPDB:GN000021; CESP:T08A11.1
A;Experimental source: clone T08A11
R;McMurray, A.
submitted to the EMBL Data Library, January 1998
A;Reference number: Z20284
A;Accession: T26904
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1872 <W12>
A;Cross-references: EMBL:AL021180; PIDN:CAA15982.1; GSPDB:GN000021; CESP:T08A11.1
A;Experimental source: clone Y44F5A
C;Genetics:
A;Gene: CESP:T08A11.1
A;Map position: 3
A;Introns: 58/3; 330/3; 366/3; 498/3; 544/3; 605/3; 896/3; 924/1; 1023/1; 1139/2; 1217/3;

Query Match 55.1%; Score 38; DB 2; Length 1872;
Best Local Similarity 50.0%; Pred. No. 3.9e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 EGGPGTTSNRLD 12
:|||||:
Db 338 KGAPGLITNRID 349

Search completed: April 28, 2005, 18:26:05
Job time: 19.875 secs

Result No.	Query	Score	Match	Length	DB	ID	Description
1	43	62.3	149	1	RNH CAUCR	Q9a341	caulobacter
2	40	58.0	229	1	GL12 ARATH	Q9m8x1	arabidopsis
3	40	58.0	344	1	MDBI RAT	P51657	rattus norv
4	38	55.1	293	1	MDCE PSEAE	Q91689	pseudomonas
5	38	55.1	293	1	SUCD MERTH	Q28663	methanobact
6	38	55.1	352	1	FLIM BORBU	Q27511	borrelia bu
7	37	53.6	222	1	AROD AERPE	Q9y8k1	oryzarium p
8	37	53.6	281	1	NPDI COREF	Q8fuc8	homo sapien
9	37	53.6	285	1	COT2 HUMAN	Q9bxj5	homo sapien
10	37	53.6	411	1	K1CI RABIT	Q28706	oryctolagus
11	37	53.6	494	1	K1CI HUMAN	Q9456	homo sapien
12	37	53.6	2796	1	CSM3 MOUSE	Q80t79	mus musculus
13	37	53.6	3670	1	CSM3 HUMAN	Q74207	homo sapien
14	36	52.2	148	1	RNH PSEAE	Q91289	pseudomonas
15	36	52.2	152	1	PAZH AUSSU	Q9pug7	austrorhiza
16	36	52.2	159	1	CDNA MOUSE	P39689	mus musculus
17	36	52.2	280	1	DLX3 ANEME	Q90229	ambystoma m
18	36	52.2	320	1	UP3B HUMAN	Q9bt76	homo sapien
19	36	52.2	330	1	RUVB CHLMU	Q9pkx8	chlamydia m
20	36	52.2	334	1	RUVB CHLTR	O84044	chlamydia t
21	36	52.2	341	1	S122 ARATH	Q9svc2	arabidopsis
22	36	52.2	379	1	AMPI SCHPO	O59730	schizosacch
23	36	52.2	417	1	CBP2 HUMAN	P48052	homo sapien
24	36	52.2	866	1	FIBA HUMAN	P04671	homo sapien
25	36	52.2	881	1	FTSK RHIME	Q92189	rhizobium m
26	36	52.2	1002	1	CLAN HUMAN	Q96jq2	homo sapien
27	36	52.2	1204	1	DNBI V2VD	P09246	varicella-z
28	36	52.2	2436	1	ABC2 HUMAN	Q9bzc7	homo sapien
29	35.5	51.4	593	1	STB2 CANEA	Q28288	canis famil
30	35.5	51.4	801	1	PHK BRAJA	Q98887	bradyrhizob
31	35	50.7	84	1	YADK MICLU	P33111	micrococcus
32	35	50.7	159	1	MPA1 CARBE	P38949	carpinus be
33	35	50.7	159	1	MPAA CORAV	Q08407	corylus ave

107 33 47.8 335 1 XYNB_STRLI
 108 33 47.8 337 1 RUVE_CHLPN
 109 33 47.8 349 1 KAPA_CANFA
 110 33 47.8 350 1 KAPA_BOVIN
 111 33 47.8 350 1 KAPA_CRIGR
 112 33 47.8 350 1 KAPA_HUMAN
 113 33 47.8 350 1 KAPA_MOUSE
 114 33 47.8 350 1 KAPA_PIG
 115 33 47.8 350 1 KAPA_RAT
 116 33 47.8 350 1 KAPA_SHEEP
 117 33 47.8 352 1 KAPC_DROME
 118 33 47.8 369 1 DCHS_RHIL0
 119 33 47.8 380 1 GRK_EACHD
 120 33 47.8 395 1 ATPB_BLEOC
 121 33 47.8 395 1 ATPB_MICPL
 122 33 47.8 400 1 PCAF_PSEPU
 123 33 47.8 412 1 BLI4_NEUCR
 124 33 47.8 413 1 AMV3_WHEAT
 125 33 47.8 413 1 E2F4_HUMAN
 126 33 47.8 420 1 YTF3_RHILT
 127 33 47.8 421 1 NUF2_RHIME
 128 33 47.8 425 1 FUTC_DROME
 129 33 47.8 439 1 AMTB_SALTY
 130 33 47.8 440 1 RHGA_ASPAC
 131 33 47.8 440 1 Y77A_STRCO
 132 33 47.8 465 1 TFSN_MOUSE
 133 33 47.8 466 1 GUNS_THEFU
 134 33 47.8 473 1 ATPB_PTECS
 135 33 47.8 473 1 ATPB_PTEES
 136 33 47.8 477 1 GTRB_HUMAN
 137 33 47.8 477 1 KCS2_HUMAN
 138 33 47.8 477 1 KCS2_MOUSE
 139 33 47.8 477 1 KCS2_RAT
 140 33 47.8 486 1 ATPB_DENPU
 141 33 47.8 490 1 ENGA_CHLTR
 142 33 47.8 492 1 GVD2_HALN1
 143 33 47.8 492 1 GVD2_HALN2
 144 33 47.8 493 1 DCMC_HUMAN
 145 33 47.8 495 1 ATPB_PTEAQ
 146 33 47.8 498 1 U713_MOUSE
 147 33 47.8 498 1 U713_RAT
 148 33 47.8 500 1 U713_HUMAN
 149 33 47.8 505 1 Y895_MYCTU
 150 33 47.8 525 1 MEPE_HUMAN

ALIGNMENTS

RESULT 1
 ID RNH_CAUCR STANDARD; PRT; 149 AA.
 AC O9A341;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DE 15-MAR-2004 (Rel. 43, Last annotation update)
 GN RNHA OR CC3365.
 OS Caulobacter crescentus.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacteriales;
 OC Caulobacteraceae; Caulobacter.
 OX NCBI_TaxID=155892;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 19089 / CB15;
 RX MEDLINE=21173698; PubMed=11259647;
 RA Nierman W.C., Feldblyum T.V., Iaub M.T., Paulsen I.T., Nelson K.E.,
 RA Eisen J.A., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
 RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
 RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
 RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
 RA Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;

RT
 RL "Complete genome sequence of *Caulobacter crescentus*.";
 Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
 CC -!- FUNCTION: This enzyme is an endonuclease that degrades the RNA of
 CC RNA-DNA hybrids specifically (By similarity).
 CC -!- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
 CC phosphonooxester.
 CC -!- COFACTOR: Binds 1 magnesium ion per subunit (By similarity).
 CC -!- SUBUNIT: Monomer (By similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
 CC -!- SIMILARITY: Belongs to the RNase H family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF005998; AAK25327.1; -;
 DR PIR; C87666; C87666.
 DR HSSP; P00647; 1RBR.
 DR TIGR; CC365; -;
 DR HAMAP; MF_00042; -; 1.
 DR InterPro; IPR002156; RNaseH.
 DR Pfam; PF00075; rnaseh; 1.
 DR KX HydroLase; Nuclease; Endonuclease; Magnesium; Complete proteome.
 KW METAL 10 10 MAGNESIUM (BY SIMILARITY).
 FT METAL 48 48 MAGNESIUM (BY SIMILARITY).
 FT METAL 70 70 MAGNESIUM (BY SIMILARITY).
 FT METAL 134 134 MAGNESIUM (BY SIMILARITY).
 SQ SEQUENCE 149 AA; 16674 MW; CB2271A33782435C CRC64;

 Query Match 62.3%; Score 43; DB 1; Length 149;
 Best Local Similarity 63.6%; Pred. No. 1.7;
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 OY 2 GGGGTTNRDL 12
 DB 38 GGGGTTNRME 48

 RESULT 2
 ID GL12_ARATH STANDARD; PRT; 229 AA.
 AC O9M8X1;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Putative germin-like protein subfamily 1 member 2 precursor.
 GN AT3G04150 OR T6K12.23.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=21016720; PubMed=11130713;
 RA Salanoubat M., Lemcke K., Rieger M., Ansoerge W., Unseld M.,
 RA Farmann B., Valle G., Bloeker H., Perez-Alonso M., Obermaier B.,
 RA Deleney M., Boutry M., Grivell L.A., Mache R., Puigdomenech P.,
 RA De Simone V., Choisme N., Artiguenave F., Robert C., Brottier P.,
 RA Wincker P., Cattolico L., Weissenbach J., Saurin W., Quetier P.,
 RA Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V.,
 RA Wurmbach E., Drzonek H., Erile H., Jordan N., Bangert S.,
 RA Wiedemann R., Kranz H., Voss H., Holland R., Brandt P., Nyakatura G.,
 RA Vezzi A., D'Angelo M., Pallavicini A., Toppi S., Simonati B.,
 RA Conrad A., Hornischer K., Kauer G., Loehner T.-H., Nordsiek G.,
 RA Reichelt J., Scharfe M., Schoen O., Bargues M., Terol J., Climent J.,
 RA Navarro P., Collado C., Perez-Perez A., Ottenwaelder B., Duchemin D.,
 RA Cooke R., Laurie M., Berger-Llauro C., Purnelle B., Masuy D.,

de Haan M., Maarse A.C., Alcaraz J.-P., Cottet A., Casacuberta E.,
 RA Monfort A., Argirou A., Flores M., Liguori R., Vitale D.,
 RA Mannhaupt G., Haase D., Schoof H., Ruda S., Zaccaria P., Mewes H.-W.,
 RA Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,
 RA Rooney T., Rizzo M., Walts A., Utterback T., Fujii C.Y., Shea T.P.,
 RA Creasy T.H., Haas B., Maiti R., Wu D., Peterson J., van Aken S.,
 RA Pai G., Miltscher J., Sellers P., Gill J.E., Feldblyum T.V.,
 RA Preuss D., Lin X., Nierman W.C., Salzberg S.L., White O., Venter J.C.,
 RA Fraser C.M., Kaneko T., Nakamura Y., Sato S., Kato T., Asamizu E.,
 RA Sasamoto S., Kimura T., Idegawa K., Kawashima K., Kishida Y.,
 RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
 RA Nakayama S., Nakazaki N., Shinpo S., Takeuchi C., Wada T.,
 RA Watanabe A., Yamada M., Yasuda M., Tabata S.;
 RT "Sequence and analysis of chromosome 3 of the plant Arabidopsis
 thaliana.";
 RL Nature 408:820-822 (2000).
 CC -!- FUNCTION: May play a role in plant defense. Has probably no
 CC oxalate oxidase activity even if the active site is conserved.
 CC -!- SUBUNIT: Oligomer (believed to be a pentamer but probably hexamer)
 CC (By similarity).
 CC -!- SIMILARITY: Apoplast (By similarity).
 CC -!- SUBCELLULAR LOCATION: Apoplast (By similarity).
 CC -!- SIMILARITY: Belongs to the germin family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AC016829; AAP26798.1; -.
 CC HSSP: P45850; 1F12.
 CC InterPro: IPR006045; Cupin.
 CC InterPro: IPR007113; Cupin_sup.
 CC InterPro: IPR001929; Germin.
 CC Pfam: PF00190; Cupin; 1.
 CC PRINTS: PR00325; GERMIN.
 CC PROSITE: PS00725; GERMIN; FALSE NEG.
 KW Apoplast; Cell wall; Signal; Glycoprotein; Manganese; Metal-binding;
 KW Multigene family; Hypothetical protein.
 FT SIGNAL 1 24 POTENTIAL.
 FT CHAIN 25 229 PUTATIVE GERMIN-LIKE PROTEIN
 FT SUBFAMILY 1 MEMBER 2.
 FT METAL 112 112 MANGANESE (BY SIMILARITY).
 FT METAL 114 114 MANGANESE (BY SIMILARITY).
 FT SITE 119 119 PROBABLE NON-FUNCTIONAL MANGANESE-BINDING
 FT SITE.
 FT METAL 163 163 MANGANESE (BY SIMILARITY).
 FT DISULFID 34 50 BY SIMILARITY.
 FT CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 71 71 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 139 139 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 229 AA; 24648 MW; D19ACDB9E93A577A CRC64;
 Query Match 58.0%; Score 40; DB 1; Length 229;
 Best Local Similarity 80.0%; Pred. No. 9.2;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 4 PGTTNRLDA 13
 DB 69 PGNTNRLGA 78
 |||||
 RESULT 3
 ID DB1 RAT STANDARD; PRT; 344 AA.
 AC P51657;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Estradiol 17 beta-dehydrogenase 1 (EC 1.1.1.62) (17-beta-HSD 1)
 DE (17-beta-hydroxysteroid dehydrogenase 1).

GN HSD17B1 OR EDH17B1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Ovary;
 RX MEDLINE=95009707; PubMed=7925110;
 RA Gherevich S., Nokelainen P., Poutanen M., Orava M.,
 RA Autio-Harminen H., Rajaniemi H., Viikio R.;
 RT "rat 17 beta-hydroxysteroid dehydrogenase type 1: primary structure
 RT and regulation of enzyme expression in rat ovary by
 RT diethylstilbestrol and gonadotropins in vivo.";
 RL Endocrinology 135:1477-1487 (1994).
 RN [2]
 RC SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Kidney;
 RX MEDLINE=96198748; PubMed=8612487;
 RA Akinola L.A., Poutanen M., Viikio R.;
 RT "Cloning of rat 17 beta-hydroxysteroid dehydrogenase type 2 and
 RT characterization of tissue distribution and catalytic activity of rat
 RT type 1 and type 2 enzymes.";
 RL Endocrinology 137:1572-1579 (1996).
 RN [3]
 RC SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Testis;
 RX MEDLINE=98201618; PubMed=9524272;
 RA Akinola L.A., Poutanen M., Pelcoketo H., Viikio R., Viikio P.;
 RT "Characterization of rat 17 beta-hydroxysteroid dehydrogenase type 1
 RT gene and mRNA transcripts.";
 RL Gene 208:229-238 (1998).
 CC -!- FUNCTION: FAVORS THE REDUCTION OF ESTROGENS AND ANDROGENS.
 CC -!- CATALYTIC ACTIVITY: Estradiol-17-beta + NAD(P) (+) = estrone +
 CC NAD(P)H.
 CC -!- PATHWAY: Estrogens biosynthesis.
 CC -!- SUBUNIT: Homodimer (By similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases
 CC (SDR) family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: X78811; CAA55389.1; -.
 CC EMBL: X97754; CAA66349.1; -.
 CC EMBL: X98038; CAA66657.1; -.
 CC PIR: S57901; S57901.
 CC HSSP: P14061; 1FDW.
 CC InterPro: IPR002198; ADH short.
 CC Pfam: PF00106; adh short; 1.
 CC PRINTS: PR00080; SDRFAMILY.
 CC PROSITE: PS00061; ADH_SHORT; 1.
 KW Steroid biosynthesis; Oxidoreductase; NAD; Multigene family.
 FT NP BIND 3 32 NAD (BY SIMILARITY).
 FT ACT_SITE 143 143 SUBSTRATE BINDING (BY SIMILARITY).
 FT ACT_SITE 156 156 BY SIMILARITY.
 SQ SEQUENCE 344 AA; 36967 MW; CBSFC139FB3995E CRC64;
 Query Match 58.0%; Score 40; DB 1; Length 344;
 Best Local Similarity 61.5%; Pred. No. 14;
 Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 1 EGGPGTTSNRLDA 13
 DB 198 EGGPGGALERADA 210
 |||||

```

RESULT 4
MDCB_PSEAE STANDARD; PRT; 293 AA.
ID MDCB_PSEAE
AC Q91689;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Probable 2-(5''-triphosphoribosyl)-3'-dephosphocoenzyme-A synthase
DE (EC 2.7.8.25) (2-(5''-triphosphoribosyl)-3'-dephospho-CoA synthase).
GN MDCB OR PA0209.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Legrou M.,
RA Garber R.L., Goltzer L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.";
RL Nature 406:959-964 (2000).
CC -!- FUNCTION: Involved in the formation of 2-(5''-phosphoribosyl)-3'-
CC dephosphocoenzyme-A, the prosthetic group of the acyl-carrier
CC protein of the malonate decarboxylase (By similarity).
CC -!- CATALYTIC ACTIVITY: ATP + 3-dephospho-CoA = 2''-(5''-
CC triphosphoribosyl)-3'-dephospho-CoA + adenine.
CC -!- SIMILARITY: Belongs to the citG/mdcB family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AB004459; AAG03598.1; -.
CC FIRM; F83618; F83618.
CC HAMAP; MF_00397; -.
CC InterPro; IPR002736; CitG.
CC Pfam; PF01874; CitG; 1.
CC Transferrase; Complete proteome.
SQ SEQUENCE 293 AA; 30533 MW; A9A085F070C680B3 CRC64;
Query Match 55.1%; Score 38; DB 1; Length 293;
Best Local Similarity 61.5%; Pred. No. 27;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 1 EGGPGTTSNRLDA 13
Db 193 EGGAGEQNRLDA 205
RESULT 5
SUCD_METHH STANDARD; PRT; 293 AA:
ID SUCD_METHH
AC O26663;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Succinyl-CoA synthetase alpha chain (EC 6.2.1.5) (SCS-alpha).
GN SUCD OR MTH563.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
OX NCBI_TaxID=187420;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Delta H;
RX MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jiwan N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RL J. Bacteriol. 179:7135-7155(1997).
CC -!- CATALYTIC ACTIVITY: ATP + succinate + CoA = ADP + succinyl-CoA +
CC phosphate.
CC -!- SIMILARITY: Belongs to the succinate/malate CoA ligase alpha
CC subunit family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AB000839; AAB85069.1; -.
CC FIRM; E69174; E69174.
CC HSSP; P07459; ISCU.
CC InterPro; IPR003781; CoA binding.
CC InterPro; IPR005810; CoA_lig_alpha.
CC InterPro; IPR005811; CoA_ligase.
CC Pfam; PF02629; CoA binding; 1.
CC Pfam; PF00549; ligase-CoA; 1.
CC PRINTS; PR01798; SCOASYNTHASE.
CC TIGRfams; TIGR01019; sucCoalalpha; 1.
CC PROSITE; PS00399; SUCCINYL_COA_LIG_2; 1.
CC PROSITE; PS01216; SUCCINYL_COA_LIG_1; 1.
CC Ligase; Tricarboxylic acid cycle; ATP-binding; Phosphorylation;
KW Complete proteome.
FT ACT_SITE 249 249 TELE-PHOSPHOHISTIDINE INTERMEDIATE
FT (BY SIMILARITY).
SQ SEQUENCE 293 AA; 30810 MW; 2E9769305CE51DA5 CRC64;
Query Match 55.1%; Score 38; DB 1; Length 293;
Best Local Similarity 53.8%; Pred. No. 27;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 1 EGGPGTTSNRLDA 13
Db 255 EGGGTASSKREA 267
RESULT 6
FLIM_BORBU STANDARD; PRT; 352 AA.
ID FLIM_BORBU
AC Q57511;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Flagellar motor switch protein flim.
GN FLIM OR BB0278.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;
RA Dunn J.U., Butler-Lofredo L., Kieleczawa J., Medalle J., Luft B.J.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=212;
RA Ge Y., Charon N.W.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=AFCC 35210 / B31;
RA MEDLINE=98065943; PubMed=9403685;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA Dougherty B., Tomb J.-P., Fleischmann R.D., Richardson D.,
RA Peterson J., Kervallage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA Utterback T., Watney L., McDonald L., Artiach P., Bowman C.,
RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi";
RL Nature 390:580-586 (1997).
CC -!- FUNCTION: FLIM IS ONE OF THREE PROTEINS (FLIG, FLIN, FLIM) THAT
CC FORM A SWITCH COMPLEX THAT IS PROPOSED TO BE LOCATED AT THE BASE
CC OF THE BASAL BODY. THIS COMPLEX INTERACTS WITH THE CHEY AND CHEZ
CC CHEMOTAXIS PROTEINS. IN ADDITION TO CONTRACTING COMPONENTS OF THE
CC MOTOR THAT DETERMINE THE DIRECTION OF FLAGELLAR ROTATION (BY
CC SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Membrane-associated (Potential).
CC -!- SIMILARITY: Belongs to the flim family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U43739; AA85602.1; -.
CC EMBL; L75945; AA858965.1; -.
CC EMBL; AB001137; AAC66670.1; -.
CC PIR; F70134; F70134.
CC TIGR; B80278; -.
CC InterPro; IPR001689; Flag_Flim.
CC ProDom; PD001543; Spoa.
CC Pfam; PF02154; Flim; 1.
CC PRINTS; PR00955; FLGMOTORFLIM.
CC ProDom; PD001777; Spoa; 1.
CC TIGRFAMs; TIGR01397; flim_switch; 1.
KW Chemotaxis; Flagellum; Flagellar rotation; Membrane;
KW Complete proteome.
SQ SEQUENCE 352 AA; 39284 MW; 2454D37BBF51C977 CRC64;

Query Match 55.1%; Score 38; DB 1; Length 352;
Best Local Similarity 72.7%; Pred. No. 33;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 GPGTTSNRLDA 13
DB 250 GVGTTSENLD 260
|||||:|||||

RESULT 7
AROD AERPE STANDARD; PRT; 222 AA.
AC Q9YEL;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 3-dehydroquinase dehydratase (EC 4.2.1.10) (3-dehydroquinase) (Type I
DE DHQase).
GN AROD OR APE0577.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
OC Desulfurococaceae; Aeropyrum.

```

```

OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=K1;
RA MEDLINE=99310339; PubMed=10382966;
RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1.";
RL DNA Res. 6:83-101 (1999).
CC -!- CATALYTIC ACTIVITY: 3-dehydroquinase = 3-dehydroshikimate + H(2)O.
CC -!- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;
CC third step.
CC -!- SIMILARITY: Belongs to the type-I 3-dehydroquinase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AP000060; BAA79545.1; -.
CC PIR; A72643; A72643.
CC HAMAP; MF_00214; -.
CC InterPro; IPR001381; DHQuinase_I.
CC PROSITE; PS01028; DEHYDROQUINASE_I; FALSE NEG.
KW Aromatic amino acid biosynthesis; Lyase; Complete proteome.
FT ACT SITE 117 117 BY SIMILARITY.
FT ACT SITE 143 143 SCHIFF-BASE INTERMEDIATE (BY SIMILARITY).
SQ SEQUENCE 222 AA; 23153 MW; A77F9F32E3303823 CRC64;

Query Match 53.6%; Score 37; DB 1; Length 222;
Best Local Similarity 63.6%; Pred. No. 30;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 GPGTTSNRLDA 13
DB 174 GPGTLASRLTA 184
|||||:|||||

RESULT 8
NPDI COREF STANDARD; PRT; 281 AA.
AC Q8FUC8;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE NAD-dependent deacetylase 1 (EC 3.5.1.-) (Regulatory protein SIR2
DE homolog 1).
GN NPDAI OR CE0092.
OS Corynebacterium efficiens.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=152794;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=YS-314 / AJ 12310 / DSM 44549 / JCM 11189;
RA MEDLINE=22723752; PubMed=12840036;
RA Nishio Y., Nakamura Y., Kawarabayashi Y., Usuda Y., Kimura E.,
RA Sugimoto S., Matsui K., Yamagishi A., Kikuchi H., Ikeo K.,
RA Gojobori T.;
RT "Comparative complete genome sequence analysis of the amino acid
RT replacements responsible for the thermostability of Corynebacterium
RT efficiens.";
RL Genome Res. 13:1572-1579 (2003).
CC -!- FUNCTION: Modulates the activities of several enzymes which are
CC inactive in their acetylated form (By similarity).

```

CC -!- CATALYTIC ACTIVITY: NAD(+) + an acetylprotein = nicotinamide + O-
 CC acetyl-ADP-ribose + a protein.
 CC -!- COFACTOR: Binds 1 zinc ion and NAD (By similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC -!- SIMILARITY: Belongs to the sirutin family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF005214; BAC16902.1; ALT_INIT.
 CC HAMAP; MF_01121; -. 1.
 CC InterPro; IPR003000; SIR2.
 CC Pfam; PF02146; SIR2; 1.
 CC PROSITE; PS03005; SIRUTIN; 1.
 CC HydroLase; NAD; Metal-binding; Zinc; Complete proteome.
 CC DEACETYLASE SIRUTIN-TYPE.
 CC FT DOMAIN 1 281
 CC FT ACT_SITE 120 120 NAD BINDING (BY SIMILARITY).
 CC FT METAL 128 128 BY SIMILARITY.
 CC FT METAL 131 131 ZINC (BY SIMILARITY).
 CC FT METAL 183 183 ZINC (BY SIMILARITY).
 CC FT METAL 186 186 ZINC (BY SIMILARITY).
 CC SQ SEQUENCE 281 AA; 30298 MW; 7EF99C3DE0012680 CRC64;
 Query Match 53.6%; Score 37; DB 1; Length 281;
 Best Local Similarity 54.5%; Pred. No. 38;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 2 GGFPGTTSNR 12
 Db 251 GGFGRADHRVD 261
 ||||| :||
 ||||| :||
 RESULT 9
 CQ22_HUMAN
 ID -CQ22_HUMAN STANDARD; PRT; 285 AA.
 AC OSBXJ5;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Complement-c1q tumor necrosis factor-related protein 2 precursor.
 GN C1QTNF2 OR CTRP2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RN SEQUENCE FROM N.A.
 RA Piddington C.S., Bishop P.;
 RT "Homo sapiens complement-c1q tumor necrosis factor-related protein.";
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RN SEQUENCE FROM N.A.
 RA Straussberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RC MEDLINE=22398257; PubMed=12477932;
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Tashiroki S., Carninci P., Prange C.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Heiton E., Kettaman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -!- SIMILARITY: Contains 1 collagenous domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF329836; AKK17960.1; -.
 CC EMBL; BC011699; AAH11699.1; -.
 CC EMBL; BC054506; AAH54506.1; -.
 CC Genew; HGNC:14325; C1QTNF2.
 CC InterPro; IPR001073; C1Q.
 CC InterPro; IPR008160; Collagen.
 CC InterPro; IPR008983; TNF_like.
 CC Pfam; PF00386; C1q; 1.
 CC Pfam; PF01391; Collagen; 2.
 CC PRINTS; PR00007; COMPLEMENTC1Q.
 CC SMART; SM01110; C1Q; 1.
 CC DR PROSITE; PS01113; C1Q; 1.
 CC Collagen; Signal.
 KW COLLAGEN, Signal.
 FT CHAIN 1 15
 FT SIGNAL 16 285
 FT DOMAIN 40 141
 FT DOMAIN 143 285
 FT SEQUENCE 285 AA; 29952 MW; 7E31FF9868D4EDFA CRC64;
 Query Match 53.6%; Score 37; DB 1; Length 285;
 Best Local Similarity 70.0%; Pred. No. 39;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 EGGPGTTSNR 10
 Db 81 EGGPGRTGGR 90
 ||||| ||||
 ||||| ||||
 RESULT 10
 K1CL_RABIT
 ID -K1CL_RABIT STANDARD; PRT; 411 AA.
 AC Q28706;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Keratin, type I cytoskeletal 12 (Cytokeratin 12) (Fragment).
 GN KRT12.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP TISSUE=Cornea;
 RC MEDLINE=94192891; PubMed=7511548;
 RA Wu R.-L., Zhu G., Galvin S., Xu C., Haseba T., Chaloin-Dufau C.,
 RA Houailly D., Wei Z.-G., Lavker R.M., Kao W.-Y., Sun T.-T.;
 RT "Lineage-specific and differentiation-dependent expression of K12
 RT keratin in rabbit corneal/limbal epithelial cells: cDNA cloning and
 RT northern blot analysis.";
 RT Differentiation 55:137-144 (1994).
 CC -!- FUNCTION: May play a unique role in maintaining the normal corneal
 CC epithelial function (By similarity).
 CC -!- SUBUNIT: Heterotetramer of two type I and two type II keratins.

Keratin 3 associates with keratin 12.
 -!- TISSUE SPECIFICITY: Cornea specific. Associated mainly with all layers of the central corneal epithelium and also found in the suprabasal limbal epithelium.
 -!- MISCELLANEOUS: There are two types of cytoskeletal and microfibrillar keratin, I (acidic) and II (neutral to basic) (40-55 and 56-70 kDa, respectively).
 -!- SIMILARITY: Belongs to the intermediate filament family.

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

 ENBL; X77665; CAA54741.1; -;
 PIR; S45318;
 InterPro; IPR001664; IF.
 DR InterPro; IPR002957; Keratin I.
 DR Pfam; PF00038; filament; 1.
 DR PRINTS; PR01248; TYPE1KERATIN.
 DR PROSITE; PS00226; IF; 1.
 KW Intermediate filament; Coiled coil; Keratin.
 FT NON_TER 1 1
 FT DOMAIN <1 42 HEAD.
 FT DOMAIN 43 354 ROD.
 FT DOMAIN 355 411 TAIL.
 FT DOMAIN 43 78 COIL 1A.
 FT DOMAIN 83 101 LINKER 1.
 FT DOMAIN 102 193 COIL 1B.
 FT DOMAIN 194 216 LINKER 12.
 FT DOMAIN 217 354 COIL 2.
 SQ SEQUENCE 411 AA; 45726 MW; 0B2E9A4D9DCC350 CRC64;
 Query Match 53.6%; Score 37; DB 1; Length 411;
 Best Local Similarity 58.3%; Pred. No. 58;
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 2 GGPQTTSNRLLDA 13
 DB 196 GGPGEVSVEMDA 207
 RESULT 11
 K1CL_HUMAN STANDARD; PRT; 494 AA.
 AC Q99456;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Keratin, type I Cytoskeletal 12 (Cyokeratin 12) (K12) (CK 12).
 GN KRT12.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Cornea;
 RX MEDLINE=96335085; PubMed=8759347;
 RA Nishida K., Adachi W., Shimizu-Matsumoto A., Kinoshita S., Mizuno K., Matsubara K., Okubo K.;
 RT "A gene expression profile of human corneal epithelium and the isolation of human keratin 12 cDNA."
 RL Invest. Ophthalmol. Vis. Sci. 37:1800-1809(1996).
 RN [2]
 RP SEQUENCE FROM N.A., AND VARIANTS MCD G1Y-135; ILE-135; ARG-140 AND ASP-429.
 RX MEDLINE=98153623; PubMed=9399908;
 RA Nishida K., Honma Y., Dota A., Kawasaki S., Adachi W., Nakamura T., Quantock A.J., Hosotani H., Yamamoto S., Okada M., Shimomura Y.,

Kinoshita S.;
 RT "Isolation and chromosomal localization of a cornea-specific human keratin 12 gene and detection of four mutations in Meesmann corneal epithelial dystrophy";
 RL Ann. J. Hum. Genet. 61:1268-1275(1997).
 RN [3]
 RP SEQUENCE FROM N.A., VARIANTS MCD THR-129 AND THR-135, AND VARIANT TRP-20.
 RX MEDLINE=20113327; PubMed=10644419;
 RA Corden L.D., Swensson O., Swensson B., Smith F.J.D., Rochels R., Utitto J., McLean W.H.I.;
 RT "Molecular genetics of Meesmann's corneal dystrophy: ancestral and novel mutations in keratin 12 (K12) and complete sequence of the human KRT12 gene";
 RL Exp. Eye Res. 70:41-49(2000).
 RN [4]
 RP VARIANTS MCD THR-135 AND LEU-143.
 RX MEDLINE=97315826; PubMed=9171831;
 RA Irvine A.D., Corden L.D., Swensson O., Swensson B., Moore J.E., Frazer D.G., Smith F.J.D., Knowlton R.G., Christophers E., Rochels R., Utitto J., McLean W.H.I.;
 RT "Mutations in cornea-specific keratin K3 or K12 genes cause Meesmann's corneal dystrophy";
 RL Nat. Genet. 16:184-187(1997).
 RN [5]
 RP VARIANT MCD PRO-130.
 RX MEDLINE=20245342; PubMed=10781519;
 RA Corden L.D., Swensson O., Swensson B., Rochels R., Wannke B., Thiel H.J., McLean W.H.I.;
 RT "A novel keratin 12 mutation in a German kindred with Meesmann's corneal dystrophy";
 RL Br. J. Ophthalmol. 84:527-530(2000).
 CC -!- FUNCTION: May play a unique role in maintaining the normal corneal epithelial function (By similarity).
 CC -!- SUBUNIT: Heterotrimer of two type I and two type II keratins. Keratin 3 associates with keratin 12.
 CC -!- TISSUE SPECIFICITY: Cornea specific.
 CC -!- DISEASE: Defects in KRT12 are a cause of juvenile epithelial corneal dystrophy of Meesmann (MCD) [MIM:122100]. MCD is an autosomal dominant disease that causes fragility of the anterior corneal epithelium. Patients are usually asymptomatic until adulthood when rupture of the corneal microcysts may cause erosions, producing clinical symptoms such as photophobia, contact lens intolerance and intermittent diminution of visual acuity. Rarely, subepithelial scarring causes irregular corneal astigmatism and permanent visual impairment. Histological examination shows a disorganized and thickened epithelium with widespread cytoplasmic vacuolation and numerous small, round, debris-laden intraepithelial cysts.
 CC -!- MISCELLANEOUS: There are two types of cytoskeletal and microfibrillar keratin, I (acidic) and II (neutral to basic) (40-55 and 56-70 kDa, respectively).
 CC -!- SIMILARITY: Belongs to the intermediate filament family.

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

 ENBL; D78367; BAA11376.1; -;
 DR ENBL; AB007119; BAA25063.1; -;
 DR ENBL; AB007115; BAA25063.1; JOINED.
 DR ENBL; AB007116; BAA25063.1; JOINED.
 DR ENBL; AB007117; BAA25063.1; JOINED.
 DR ENBL; AB007118; BAA25063.1; JOINED.
 DR ENBL; AF137286; AAF61432.1; -;
 DR Genew; HGNC:6414; KRT12.
 DR MIM; 601687; -;
 DR MIM; 122100; -;
 DR GO; GO:0007601; P:vision; TAS.

```

DR InterPro: IPR001664; IF.
DR InterPro: IPR002957; Keratin_1.
DR Pfam: PF00038; Filament; 1.
DR PRINTS: PRO1248; TYPE1KERATIN.
DR PROSITE: PS00226; IF; 1.
KW Intermediate filament; Coiled coil; Keratin; Disease mutation.
FT DOMAIN 1 124 HEAD.
FT DOMAIN 125 435 ROD.
FT DOMAIN 126 494 TAIL.
FT DOMAIN 127 160 COIL 1A.
FT DOMAIN 128 182 COIL 1B.
FT DOMAIN 129 182 COIL 1B.
FT DOMAIN 130 274 COIL 1B.
FT DOMAIN 131 274 COIL 1B.
FT DOMAIN 132 274 COIL 1B.
FT DOMAIN 133 274 COIL 1B.
FT DOMAIN 134 274 COIL 1B.
FT DOMAIN 135 274 COIL 1B.
FT DOMAIN 136 274 COIL 1B.
FT DOMAIN 137 274 COIL 1B.
FT DOMAIN 138 274 COIL 1B.
FT DOMAIN 139 274 COIL 1B.
FT DOMAIN 140 274 COIL 1B.
FT DOMAIN 141 274 COIL 1B.
FT DOMAIN 142 274 COIL 1B.
FT DOMAIN 143 274 COIL 1B.
FT DOMAIN 144 274 COIL 1B.
FT DOMAIN 145 274 COIL 1B.
FT DOMAIN 146 274 COIL 1B.
FT DOMAIN 147 274 COIL 1B.
FT DOMAIN 148 274 COIL 1B.
FT DOMAIN 149 274 COIL 1B.
FT DOMAIN 150 274 COIL 1B.
FT DOMAIN 151 274 COIL 1B.
FT DOMAIN 152 274 COIL 1B.
FT DOMAIN 153 274 COIL 1B.
FT DOMAIN 154 274 COIL 1B.
FT DOMAIN 155 274 COIL 1B.
FT DOMAIN 156 274 COIL 1B.
FT DOMAIN 157 274 COIL 1B.
FT DOMAIN 158 274 COIL 1B.
FT DOMAIN 159 274 COIL 1B.
FT DOMAIN 160 274 COIL 1B.
FT DOMAIN 161 274 COIL 1B.
FT DOMAIN 162 274 COIL 1B.
FT DOMAIN 163 274 COIL 1B.
FT DOMAIN 164 274 COIL 1B.
FT DOMAIN 165 274 COIL 1B.
FT DOMAIN 166 274 COIL 1B.
FT DOMAIN 167 274 COIL 1B.
FT DOMAIN 168 274 COIL 1B.
FT DOMAIN 169 274 COIL 1B.
FT DOMAIN 170 274 COIL 1B.
FT DOMAIN 171 274 COIL 1B.
FT DOMAIN 172 274 COIL 1B.
FT DOMAIN 173 274 COIL 1B.
FT DOMAIN 174 274 COIL 1B.
FT DOMAIN 175 274 COIL 1B.
FT DOMAIN 176 274 COIL 1B.
FT DOMAIN 177 274 COIL 1B.
FT DOMAIN 178 274 COIL 1B.
FT DOMAIN 179 274 COIL 1B.
FT DOMAIN 180 274 COIL 1B.
FT DOMAIN 181 274 COIL 1B.
FT DOMAIN 182 274 COIL 1B.
FT DOMAIN 183 274 COIL 1B.
FT DOMAIN 184 274 COIL 1B.
FT DOMAIN 185 274 COIL 1B.
FT DOMAIN 186 274 COIL 1B.
FT DOMAIN 187 274 COIL 1B.
FT DOMAIN 188 274 COIL 1B.
FT DOMAIN 189 274 COIL 1B.
FT DOMAIN 190 274 COIL 1B.
FT DOMAIN 191 274 COIL 1B.
FT DOMAIN 192 274 COIL 1B.
FT DOMAIN 193 274 COIL 1B.
FT DOMAIN 194 274 COIL 1B.
FT DOMAIN 195 274 COIL 1B.
FT DOMAIN 196 274 COIL 1B.
FT DOMAIN 197 274 COIL 1B.
FT DOMAIN 198 274 COIL 1B.
FT DOMAIN 199 274 COIL 1B.
FT DOMAIN 200 274 COIL 1B.
FT VARIANT 1 124 HEAD.
FT VARIANT 125 435 ROD.
FT VARIANT 126 494 TAIL.
FT VARIANT 127 160 COIL 1A.
FT VARIANT 128 182 COIL 1B.
FT VARIANT 129 182 COIL 1B.
FT VARIANT 130 274 COIL 1B.
FT VARIANT 131 274 COIL 1B.
FT VARIANT 132 274 COIL 1B.
FT VARIANT 133 274 COIL 1B.
FT VARIANT 134 274 COIL 1B.
FT VARIANT 135 274 COIL 1B.
FT VARIANT 136 274 COIL 1B.
FT VARIANT 137 274 COIL 1B.
FT VARIANT 138 274 COIL 1B.
FT VARIANT 139 274 COIL 1B.
FT VARIANT 140 274 COIL 1B.
FT VARIANT 141 274 COIL 1B.
FT VARIANT 142 274 COIL 1B.
FT VARIANT 143 274 COIL 1B.
FT VARIANT 144 274 COIL 1B.
FT VARIANT 145 274 COIL 1B.
FT VARIANT 146 274 COIL 1B.
FT VARIANT 147 274 COIL 1B.
FT VARIANT 148 274 COIL 1B.
FT VARIANT 149 274 COIL 1B.
FT VARIANT 150 274 COIL 1B.
FT VARIANT 151 274 COIL 1B.
FT VARIANT 152 274 COIL 1B.
FT VARIANT 153 274 COIL 1B.
FT VARIANT 154 274 COIL 1B.
FT VARIANT 155 274 COIL 1B.
FT VARIANT 156 274 COIL 1B.
FT VARIANT 157 274 COIL 1B.
FT VARIANT 158 274 COIL 1B.
FT VARIANT 159 274 COIL 1B.
FT VARIANT 160 274 COIL 1B.
FT VARIANT 161 274 COIL 1B.
FT VARIANT 162 274 COIL 1B.
FT VARIANT 163 274 COIL 1B.
FT VARIANT 164 274 COIL 1B.
FT VARIANT 165 274 COIL 1B.
FT VARIANT 166 274 COIL 1B.
FT VARIANT 167 274 COIL 1B.
FT VARIANT 168 274 COIL 1B.
FT VARIANT 169 274 COIL 1B.
FT VARIANT 170 274 COIL 1B.
FT VARIANT 171 274 COIL 1B.
FT VARIANT 172 274 COIL 1B.
FT VARIANT 173 274 COIL 1B.
FT VARIANT 174 274 COIL 1B.
FT VARIANT 175 274 COIL 1B.
FT VARIANT 176 274 COIL 1B.
FT VARIANT 177 274 COIL 1B.
FT VARIANT 178 274 COIL 1B.
FT VARIANT 179 274 COIL 1B.
FT VARIANT 180 274 COIL 1B.
FT VARIANT 181 274 COIL 1B.
FT VARIANT 182 274 COIL 1B.
FT VARIANT 183 274 COIL 1B.
FT VARIANT 184 274 COIL 1B.
FT VARIANT 185 274 COIL 1B.
FT VARIANT 186 274 COIL 1B.
FT VARIANT 187 274 COIL 1B.
FT VARIANT 188 274 COIL 1B.
FT VARIANT 189 274 COIL 1B.
FT VARIANT 190 274 COIL 1B.
FT VARIANT 191 274 COIL 1B.
FT VARIANT 192 274 COIL 1B.
FT VARIANT 193 274 COIL 1B.
FT VARIANT 194 274 COIL 1B.
FT VARIANT 195 274 COIL 1B.
FT VARIANT 196 274 COIL 1B.
FT VARIANT 197 274 COIL 1B.
FT VARIANT 198 274 COIL 1B.
FT VARIANT 199 274 COIL 1B.
FT VARIANT 200 274 COIL 1B.
SQ SEQUENCE 494 AA; 53511 MW; 75C981380532B682 CRC64;

Query Match 53.6%; Score 37; DB 1; Length 494;
Best Local Similarity 58.3%; Pred. No. 71;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 GPGGTTSNRLDA 13
DB 277 GPGGTVSVMEDA 288

RESULT 12
CSM3_MOUSE STANDARD; PRT; 2796 AA.
AC Q80T79; Q8BVJ0; Q9D588;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE CUB and sushi multiple domains protein 3 (fragment).
GN CSM03 OR KIAA1894.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC MEDLINE=22579291; PubMed=12693553;
RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Aizawa H., Yuasa S.,
RA Nakajima D., Nagase T., Ohara O., Koga H.;
RT "Prediction of the coding sequences of mouse homologues of KIAA gene:
RT cDNAs identified by screening of terminal sequences of cDNA clones
RT randomly sampled from size-fractionated libraries.";
RL DNA Res. 10:35-48(2003).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2), AND SEQUENCE OF 2379-2796 FROM N.A.
RP (ISOFORM 1).
RC STRAIN=C57BL/6J; TISSUE=Medulla oblongata, and Testis;
RX MEDLINE=22354683; PubMed=12466851;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,

```

```

RA Nikaido I., Oosato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojibori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusci V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmer S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Sempie C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wyshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoID=Q80T79-1; Sequences=Displayed;
CC Name=2;
CC IsoID=Q80T79-2; Sequences=VSP_009053, VSP_009054, VSP_009055,
CC VSP_009056;
CC Note=No experimental confirmation available;
CC -!- SIMILARITY: Belongs to the CSM0 family.
CC -!- SIMILARITY: Contains at least 10 CUB domains.
CC -!- SIMILARITY: Contains at least 25 Sushi (SCR) domains.
CC -!- CAUTION: Ref.2 (BAC37116) sequence differs from that shown due to
CC multiple frameshifts from position 2763.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AK122567; BAC65849.1; -
CC EMBL: AK015672; BAB29924.1; -
CC EMBL: AK078076; BAC37116.1; ALT FRAME.
CC MGD: MGI:1922193; 4930500N14Rik.
CC InterPro: IPR000859; CUB.
CC InterPro: IPR000436; Sushi_SCR_CCP.
CC Pfam: PF00431; CUB; 9.
CC Pfam: PF00084; sushi; 25.
CC SMART: SM00032; CCP; 25.
CC SMART: SM00042; CUB; 9.
CC PROSITE: PS01180; CUB; 10.
CC Repeat; Transmembrane; Sushi; Alternative splicing.
KW NON TER 1 1
FT DOMAIN 1 50 SUSHI 1.
FT DOMAIN 54 162 CUB 1.
FT DOMAIN 169 222 SUSHI 2.
FT DOMAIN 226 336 CUB 2.
FT DOMAIN 341 396 SUSHI 3.
FT DOMAIN 400 508 CUB 3.
FT DOMAIN 513 569 SUSHI 4.
FT DOMAIN 573 682 CUB 4.
FT DOMAIN 687 743 SUSHI 5.

```


RESULT 14
 RNH_PSEAE STANDARD; PRT; 148 AA.
 AC Q91259;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DE 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Ribonuclease HI (EC 3.1.26.4) (RNase HI).
 GN RNHA OR PA1815.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PA01;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warriner P.,
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Goltner L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
 RA Brody L.L., Coulter S.H., Folger K.R., Kas A., Larbig K., Lim R.M.,
 RA Smith K.A., Spencer D.N., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Raizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 RT opportunistic pathogen.";
 RL Nature 406:959-964 (2000).
 CC -1- FUNCTION: This enzyme is an endonuclease that degrades the RNA of
 CC RNA-DNA hybrids specifically (By similarity).
 CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
 CC phosphonate.
 CC -1- COFACTOR: Binds 1 magnesium ion per subunit (By similarity).
 CC -1- SUBUNIT: Monomer (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (potential).
 CC -1- SIMILARITY: Belongs to the RNase H family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; A5004608; AAG05204.1; -
 CC FIR; A83418; A83418.
 CC DR HSP; P00647; LRBS.
 CC DR HAMAP; MF 00042; -; 1.
 CC DR InterPro; IPR002156; RNaseH.
 CC DR Pfam; PF00075; RNaseH; 1.
 CC KW Hydrolyase; Nuclease; Endonuclease; Magnesium; Complete proteome.
 FT METAL 12 12 MAGNESIUM (BY SIMILARITY).
 FT METAL 50 50 MAGNESIUM (BY SIMILARITY).
 FT METAL 72 72 MAGNESIUM (BY SIMILARITY).
 FT METAL 136 136 MAGNESIUM (BY SIMILARITY).
 SQ SEQUENCE 148 AA; 16697 MW; 1171AF5A9D267CD8 CRC64;
 Query Match 52.2%; Score 36; DB 1; Length 148;
 Best Local Similarity 54.5%; Pred. No. 28;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 2 GGPFTTSNRD 12
 DB 40 GEPDTNNRME 50
 RESULT 15
 PA2H AUSSU STANDARD; PRT; 152 AA.
 AC Q9PUG7;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Phospholipase A2 isozyme S17-58 precursor (EC 3.1.1.4)
 DE (Phosphatidylcholine 2-acylhydrolase) (ASPLA17).
 OS Austrelaps superbus (Australian copperhead).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
 OC Elapidae; Acanthophiinae; Austrelaps.
 OX NCBI_TaxID=29156;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Venom gland;
 RX MEDLINE=20166950; PubMed=10700385;
 RA Singh S.B., Armugam A., Kini R.M., Jayaseelan K.;
 RT "Phospholipase A(2) with platelet aggregation inhibitor activity from
 RT Austrelaps superbus venom: protein purification and cDNA cloning.";
 RL Arch. Biochem. Biophys. 375:289-303 (2000).
 CC -1- FUNCTION: PA2 catalyzes the calcium-dependent hydrolysis of the 2-
 CC acyl groups in 3-sn-phosphoglycerides. Inhibits collagen-induced
 CC platelet aggregation (By similarity).
 CC -1- CATALYTIC ACTIVITY: Phosphatidylcholine + H(2)O = 1-
 CC acylglycerophosphocholine + a fatty acid anion.
 CC -1- COFACTOR: Calcium (Probable).
 CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
 CC -1- SIMILARITY: Belongs to the phospholipase A2 family. Group I
 CC subfamily.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF184143; AAD56410.1; -
 CC DR HSP; P00593; 4BP2.
 CC DR InterPro; IPR001211; PhospholipaseA2.
 CC DR Pfam; PF00068; phoslip; 1.
 CC DR PRINTS; PR00389; PHPLIPASEA2.
 CC DR ProDom; PD000303; PhospholipaseA2; 1.
 CC DR SMART; SM00085; PA2C; 1.
 CC DR PROSITE; PS00119; PA2 ASP; 1.
 CC DR PROSITE; PS00118; PA2 HIS; 1.
 CC KW Lipid degradation; Hydrolase; Toxin; Signal; Calcium;
 CC Multigene family.
 FT SIGNAL 1 19 POTENTIAL.
 FT PROPEP 20 27 POTENTIAL.
 FT CHAIN 28 152 PHOSPHOLIPASE A2 ISOZYME S17-58.
 FT ACT_SITE 75 75 BY SIMILARITY.
 FT ACT_SITE 126 126 BY SIMILARITY.
 FT DISULFID 38 104 BY SIMILARITY.
 FT DISULFID 54 151 BY SIMILARITY.
 FT DISULFID 71 132 BY SIMILARITY.
 FT DISULFID 78 125 BY SIMILARITY.
 FT DISULFID 88 118 BY SIMILARITY.
 FT DISULFID 111 133 BY SIMILARITY.
 FT METAL 55 55 CALCIUM (VIA CARBONYL OXYGEN)
 FT METAL 57 57 (BY SIMILARITY).
 FT METAL 59 59 CALCIUM (VIA CARBONYL OXYGEN)
 FT METAL 59 59 (BY SIMILARITY).
 FT METAL 76 76 CALCIUM (VIA CARBONYL OXYGEN)
 FT METAL 76 76 (BY SIMILARITY).
 SQ SEQUENCE 152 AA; 16641 MW; 0F2BD5BDDCF04FCB CRC64;
 Query Match 52.2%; Score 36; DB 1; Length 152;
 Best Local Similarity 63.6%; Pred. No. 29;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 2 GGPFTTSNRD 12
 DB 59 GSGTTPSDELD 69

RESULT 16
CDNA_MOUSE STANDARD; PRT; 159 AA.
AC F39689;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cyclin-dependent kinase inhibitor 1 (p21) (CDK-interacting protein 1)
DE (Melanoma differentiation associated protein).
GN CDKN1A OR CIP1 OR WAF1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BXSB; TISSUE=Spleen;
RX MEDLINE=94366751; PubMed=8084607;
RA Huppi K., Siwarski D., Dosik J., Michieli P., Chedid M., Reed S.,
RA Mock B., Givol D., Mushinski J.F.,
RT "Molecular cloning, sequencing, chromosomal localization and
RT expression of mouse p21 (Waf1).";
RL Oncogene 9:3017-3020(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95316868; PubMed=7796420;
RA El-Deiry W.S., Tokino T., Waldman T., Velculescu V., Oliner J.D.,
RA Burrell M., Hill D.E., Res J.L., Hamilton S.R., Kinzler K.W.,
RA Vogelstein B.;
RT "Topological control of p21WAF1/CIP1 expression in normal and
RT neoplastic tissues.";
RL Cancer Res. 55:2910-2919(1995).
RN [3]
RP SEQUENCE OF 1-143 FROM N.A.
RX MEDLINE=94061997; PubMed=8242752;
RA El-Deiry W.S., Tokino T., Velculescu V.E., Levy D.B., Parsons R.,
RA Trent J.M., Lin D., Mercer W.E., Kinzler K.W., Vogelstein B.;
RT "WAF1, a potential mediator of p53 tumor suppression.";
RL Cell 75:817-825(1993).
CC -!- FUNCTION: May be the important intermediate by which p53 mediates
CC its role as an inhibitor of cellular proliferation in response to
CC DNA damage. May bind to and inhibit cyclin-dependent kinase
CC activity, preventing phosphorylation of critical cyclin-dependent
CC kinase substrates and blocking cell cycle progression.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- INDUCTION: By p53, mezerein (antileukemic compound) and interferon
CC beta.
CC -!- SIMILARITY: Belongs to the CDI family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U09507; AAB60456.1; -;
CC FIR; U24173; AAC52220.1; -;
CC FIR; A49438; A49438.
CC MGD; MGI:104556; Cdkn1a.
CC InterPro; IPR003175; CDI.
CC Pfam; PF02234; CDI.1
KW Protein kinase inhibitor; Cell cycle; Nuclear protein; Zinc-finger.
FT ZN FING 12 40 C4-TYPE (POTENTIAL).
FT CONFLICT 30 30 R -> S (IN REF. 3).
FT CONFLICT 56 57 TP -> RQ (IN REF. 3).
SQ SEQUENCE 159 AA; 17785 MW; 37B7C22B9A2FD089 CRC64;
Query Match 52.2%; Score 36; DB 1; Length 159;
Best Local Similarity 66.7%; Pred. No. 31;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 2 GGGGTTSNR 10
DB 127 GGGTSGQR 135
RESULT 17
DLX3_AMBME STANDARD; PRT; 280 AA.
AC Q90229;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Homeobox protein DLX-3.
GN DLX-3.
OS Ambystoma mexicanum (Axolotl).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Caudata; Salamandroidea; Ambystomatidae;
OC Ambystoma.
OX NCBI_TaxID=8296;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97108743; PubMed=8951064;
RA Mullen L.M., Bryant S.V., Torok M.A., Blumberg B., Gardiner D.M.;
RT "Nerve dependency of regeneration: the role of Distal-less and FGF
RT signaling in amphibian limb regeneration.";
RL Development 122:3487-3497(1996).
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- SIMILARITY: Belongs to the distal-less homeobox family.
CC -!- SIMILARITY: Contains 1 homeobox domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U59480; AAB49668.1; -;
CC HSSP; P02836; 3HDD.
CC InterPro; IPR001356; Homeobox.
CC InterPro; IPR000047; HTH lambdarepressr.
CC Pfam; PF00046; homeobox.1.
CC PRINTS; PR00024; HOMEBOX.
CC PRINTS; PR00031; HTHREPRESSR.
CC ProDom; PD000010; Homeobox.1.
CC SMART; SM00389; HOX.1.
CC PROSITE; PS00027; HOMEBOX 1; 1.
CC PROSITE; PS00071; HOMEBOX 2; 1.
KW Homeobox; DNA-binding; Developmental protein; Nuclear protein.
FT DNA BIND 126 135 HOMEBOX.
SQ SEQUENCE 280 AA; 31200 MW; 0F8183097ABAC791 CRC64;
Query Match 52.2%; Score 36; DB 1; Length 280;
Best Local Similarity 63.6%; Pred. No. 57;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 2 GGGGTTSNR 12
DB 72 GGGTYSKSD 82
RESULT 18
UP3B_HUMAN STANDARD; PRT; 320 AA.
AC Q9BT76; Q86W06;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Uroplakin 3B precursor (Uroplakin IIIB) (UPIIIB) (p35).
GN UPK3B.

```

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., SUBUNIT, AND SUBCELLULAR LOCATION.
RX MEDLINE=22336509; PubMed=12446744;
RA Deng F.-W., Liang F.-X., Tu L., Resing K.A., Hu P., Supino M.,
RA Hu C.-C.A., Zhou G., Ding M., Kreibich G., Sun T.-T.;
RT "Uroplakin IIb, a urothelial differentiation marker, dimerizes with
RT uroplakin Ib as an early step of urothelial plaque assembly.";
RL J. Cell Biol. 159:685-694(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Uterus;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M.J., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smalhus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Component of the asymmetric unit membrane (AUM); a
CC highly specialized biomembrane elaborated by terminally
CC differentiated urothelial cells. May play an important role in
CC AUM-cytoskeleton interaction in terminally differentiated
CC urothelial cells. It also contributes to the formation of
CC urothelial glycocalyx which may play an important role in
CC preventing bacterial adherence (By similarity).
CC -!- SUBUNIT: Forms heterodimer with Uroplakin Ib.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC Heterodimer formation with Uroplakin Ib is a prerequisite to exit
CC out of the endoplasmic reticulum (ER).
CC -!- CAUTION: Ref.1 sequence differs from that shown due to a
CC frameshift in position 210.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AY233462; AAC9507.1; ALT_FRAME.
CC EMBL; BC004304; AA04304.1; -.
CC Genew; HGNC:21444; UPK3B.
KW Endoplasmic reticulum; Signal; Transmembrane; Glycoprotein.
FT SIGNAL 1 29
FT CHAIN 30 320
FT DOMAIN 241 266
FT TRANSMEM 241 266
FT DOMAIN 267 320
FT CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 133 133
FT CONFLICT 129 83
FT MISSING (IN REF. 1). (POTENTIAL).
SQ SEQUENCE 320 AA; 33851 MW; 64AC6E268AB0B0EB CRC64;

Query Match 52.2%; Score 36; DB 1; Length 320;
Best Local Similarity 60.0%; Pred. No. 66;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., SUBUNIT, AND SUBCELLULAR LOCATION.
RX MEDLINE=22336509; PubMed=12446744;
RA Deng F.-W., Liang F.-X., Tu L., Resing K.A., Hu P., Supino M.,
RA Hu C.-C.A., Zhou G., Ding M., Kreibich G., Sun T.-T.;
RT "Uroplakin IIb, a urothelial differentiation marker, dimerizes with
RT uroplakin Ib as an early step of urothelial plaque assembly.";
RL J. Cell Biol. 159:685-694(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Uterus;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M.J., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smalhus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Component of the asymmetric unit membrane (AUM); a
CC highly specialized biomembrane elaborated by terminally
CC differentiated urothelial cells. May play an important role in
CC AUM-cytoskeleton interaction in terminally differentiated
CC urothelial cells. It also contributes to the formation of
CC urothelial glycocalyx which may play an important role in
CC preventing bacterial adherence (By similarity).
CC -!- SUBUNIT: Forms heterodimer with Uroplakin Ib.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC Heterodimer formation with Uroplakin Ib is a prerequisite to exit
CC out of the endoplasmic reticulum (ER).
CC -!- CAUTION: Ref.1 sequence differs from that shown due to a
CC frameshift in position 210.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AY233462; AAC9507.1; ALT_FRAME.
CC EMBL; BC004304; AA04304.1; -.
CC Genew; HGNC:21444; UPK3B.
KW Endoplasmic reticulum; Signal; Transmembrane; Glycoprotein.
FT SIGNAL 1 29
FT CHAIN 30 320
FT DOMAIN 241 266
FT TRANSMEM 241 266
FT DOMAIN 267 320
FT CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 133 133
FT CONFLICT 129 83
FT MISSING (IN REF. 1). (POTENTIAL).
SQ SEQUENCE 320 AA; 33851 MW; 64AC6E268AB0B0EB CRC64;

Query Match 52.2%; Score 36; DB 1; Length 320;
Best Local Similarity 60.0%; Pred. No. 66;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 2 GGGPTTSNRL 11
DB 260 GCGGAADRL 269
||| :||
||| :||

RESULT 19
RUVB CHLMU
ID RUVB CHLMU STANDARD; PRT; 330 AA.
AC Q9PKZ8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Holliday junction DNA helicase ruvB.
GN RUVB OR TC0310.
OS Chlamydia muridarum.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83560;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MoPn / Nigg;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RA "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RA pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
CC -!- FUNCTION: The ruvA-ruvB complex in the presence of ATP renatures
CC cruciform structure in supercoiled DNA with palindromic sequence,
CC indicating that it may promote strand exchange reactions in
CC homologous recombination. RuvAB is an helicase that mediates the
CC Holliday junction migration by localized denaturation and
CC reannealing (By similarity).
CC -!- SUBUNIT: Forms a complex with ruvA (By similarity).
CC -!- SIMILARITY: Belongs to the ruvB family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF002298; AAF39175.1; -.
CC EMBL; E81717; E81717.
CC TIGR; TC0310; -.
CC HAMAP; MF_00016; -.
CC InterPro; IPR003593; AAA_ATPase.
CC InterPro; IPR003959; AAA_ATPase_cent.
CC InterPro; IPR004605; RuvB.
CC InterPro; IPR008823; RuvB_C.
CC InterPro; IPR008824; RuvB_N.
CC Pfam; PF00004; AAA; 1.
CC Pfam; PF05491; RuvB_C; 1.
CC Pfam; PF05496; RuvB_N; 1.
CC SMART; SM00382; AAA; 1.
CC TIGRfam; TIGR00635; ruvB; 1.
CC DNA repair; SOS response; ATP-binding; DNA recombination; Helicase;
KW Complete proteome.
FT NP_BIND 57 64
FT ATP (POTENTIAL).
SQ SEQUENCE 330 AA; 36875 MW; B734E5AEBC04C67 CRC64;

Query Match 52.2%; Score 36; DB 1; Length 330;
Best Local Similarity 58.3%; Pred. No. 69;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 1 EGGPGTTSNRLD 12
: ||| | |||

```

Db 135 DSGFGARSVRLD 146

RESULT 20

RUVB_CHLTR

ID_RUVB_CHLTR STANDARD; PRT; 334 AA.

AC 084044;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Holliiday junction DNA helicase ruvB.

GN RUVB OR CT040.

OS Chlamydia trachomatis.

OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.

OX NCBI_TaxID=813;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=D/UW-3/Cx;

RX MEDLINE=9900809; PubMed=9784136;

RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L., Mitchell W.P., Olinger L., Tatusev R.L., Zhao Q., Koonin E.V., Davis R.W.;

RA "Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis.";

RT Science 282:754-759(1998).

RL -1- FUNCTION: The ruvA-ruvB complex in the presence of ATP renatures cruciform structure in supercoiled DNA with palindromic sequence, indicating that it may promote strand exchange reactions in homologous recombination. RuvAB is an helicase that mediates the Holliday junction migration by localized denaturation and reannealing (By similarity).

CC -1- SUBUNIT: Forms a complex with ruvA (By similarity).

CC -1- SIMILARITY: Belongs to the ruvB family.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AE001278; AAC67630.1; -.

DR PIR; C71565; C71565.

DR HAMAP; MF 00016; - 1.

DR InterPro; IPR003593; AAA_ATPase.

DR InterPro; IPR003959; AAA_ATPase_cent.

DR InterPro; IPR004605; RuvB.

DR InterPro; IPR008823; RuvB_C.

DR InterPro; IPR008824; RuvB_N.

DR Pfam; PF00004; AAA; 1.

DR Pfam; PF05491; RuvB_C; 1.

DR Pfam; PF05496; RuvB_N; 1.

DR SMART; SM00382; AAA; 1.

DR TIGRfams; TIGR00635; ruvB; 1.

KW DNA repair; SOS response; ATP-binding; DNA recombination; Helicase;

KW Complete proteome.

FT NP_BIND 57 64 ATP (POTENTIAL).

SQ SEQUENCE 334 AA; 37289 MW; 2FF0F174C5C6B705 CRC64;

Query Match 52.2%; Score 36; DB 1; Length 334;

Best Local Similarity 58.3%; Pred. No. 69;

Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 EGGPGTTSNRLD 12

: ||| |||

Db 135 DSGFGARSVRLD 146

RESULT 21

S122 ARATH

ID S122 ARATH STANDARD; PRT; 341 AA.

AC Q9SVC2; Q9SUJ2;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Syntaxin 122 (AtSYP122) (Synt4).

GN SYP122 OR AT3G52400 OR P2206 220.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.

OX NCBI_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RA Stracke R., Palme K.;

RT "Signal peptide selection derived cDNAs from Arabidopsis thaliana leaves and guard cells.";

RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. Columbia;

RX MEDLINE=21016720; PubMed=11130713;

RA Salanoubat M., Lemcke K., Rieger M., Ansoerge W., Unseid M., Fartmann B., Valle G., Bloeker H., Perez-Alonso M., Obermaier B., Delenoy M., Boutry M., Grivell L.A., Mache R., Puigdomenech P., De Simone V., Choisme N., Artiguenave F., Robert C., Brottier P., Wincker P., Cattolico L., Weissbach J., Saurin W., Quetier F., Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Henes V., Wiedemann E., Drzonek H., Efrle H., Jordan N., Bangert S., Vezzi A., D'Angelo M., Pallavicini A., Toppo S., Simonati B., Reichelt J., Scharfe K., Kauer G., Loehert T.H., Nordsiek G., Navarro P., Collado C., Perez-Perz A., Ottenwelder B., Duchemin D., Cooke R., Laudie M., Berger-Llauro C., Purnelle B., Masuy D., De Haan M., Maarse A.C., Alcaraz J.-P., Cottet A., Casacuberta E., Monfort A., Argiriou A., Flores M., Liguori R., Vitale D., Mayer K.F.X., Haase D., Schoof H., Rudd S., Zaccaria P., Jenkins J., Rooney T., Rizzo M., Walts A., Utterback T., Fujii C.Y., Shen S., Creasy T.H., Haas B., Maiti R., Wu D., Peterson J., Van Aken S., Pai G., Militscher J., Sellers P., Gill J.E., Feildlyum T.V., Preuss D., Lin X., Nierman W.C., Salzberg S.L., White O., Venter J.C., Fraser C.M., Kaneko T., Nakamura Y., Sato S., Kato T., Asamizu E., Sasamoto S., Kimura T., Idesawa K., Kawashima K., Kishida Y., Kiyokawa S., Kohara M., Matsumoto M., Matsuno A., Muraki A., Nakayama S., Nakazaki N., Shinpo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M., Tabata S.;

RT "Sequence and analysis of chromosome 3 of the plant Arabidopsis thaliana.";

RL Nature 408:820-822(2000).

RN [4]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. Columbia;

RX MEDLINE=22954850; PubMed=14593172;

RA Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J., Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F., Kiralin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G., Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J., Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y., Arakawa T., Bann J., Banno F., Bowser L., Brooks S.Y., Carninci P., Chao Q., Choy N., Enju A., Goldsmith A.D., Gurjal M., Hansen N.F., Hayashizaki Y., Johnson-Hopson C., Hsuan V.W., Iida K., Karnes M., Khan S., Koesema E., Ishida J., Jiang P.X., Jones T., Kawai J., Kamiya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T., Satou M., Tamse R., Vaysberg M., Wallender E.K., Wong C., Yamamura Y., Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;

RT "Empirical analysis of transcriptional activity in the Arabidopsis

```

genome.";
Science 302:842-846(2003).
[5]
SEQUENCE FROM N.A.
RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.A.;
RA "Full-length cDNA from Arabidopsis thaliana.";
RA Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -i- FUNCTION: Vesicle trafficking protein that functions in the
CC secretory pathway (By similarity).
CC -i- SUBUNIT: Part of the t-SNARE complex (By similarity).
CC -i- SUBCELLULAR LOCATION: Type IV membrane protein (By similarity).
CC -i- SIMILARITY: Belongs to the syntaxin/epimorphin family.
CC -i- SIMILARITY: Contains 1 t-SNARE coiled-coil homology domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AF083808; AAM60366.1; -
CC EMBL: AJ245407; CAB52174.1; -
CC EMBL: AL050300; CAB43444.1; -
CC EMBL: AY050907; AAK93584.1; -
CC EMBL: AY091410; AAM14349.1; -
CC EMBL: AY087633; AAM65172.1; -
CC PIR: T08459; T08459.
CC PIR: T48847; T48847.
CC InterPro: IPR06012; Syntaxin.
CC InterPro: IPR06011; Syntaxin_N.
CC InterPro: IPR000727; T-SNARE.
CC Pfam: PF05739; SNARE; 1.
CC SMART: SM00804; Syntaxin; 1.
CC SMART: SM00503; SynN; 1.
CC SMART: SM00397; t-SNARE; 1.
CC PROSITE: PS00914; SYNTAXIN; 1.
CC PROSITE: PS0192; T-SNARE; 1.
CC Transport; Protein transport; Transmembrane; Coiled coil;
KW Multigene family.
FT DOMAIN 1 284
FT TRANSMEM 285 305 CYTOPLASMIC (POTENTIAL).
FT ANCHOR FOR TYPE IV MEMBRANE PROTEIN
FT (POTENTIAL).
FT DOMAIN 306 341 VESICULAR (POTENTIAL).
FT DOMAIN 64 185 COILED COIL (POTENTIAL).
FT DOMAIN 213 275 T-SNARE COILED-COIL HOMOLGY.
FT CONFLICT 214 214 N -> S (IN REF. 2 AND 5).
FT SEQUENCE 341 AA; 37837 MW; 65817034D954ABFF CRC64;
Query Match 52.2%; Score 36; DB 1; Length 341;
Best Local Similarity 60.0%; Pred. No. 71;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 EGGPGTTSNR 10
Db 123 EGGPGSSDR 132
RESULT 22
ID AMP1_SCHPO STANDARD; PRT; 379 AA.
AC O59730;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Probable methionine aminopeptidase 1 (EC 3.4.11.18) (MetAP 1)
DE (MAP 1) (Peptidase M 1).
GN SPBC3E7.10.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;

```

```

Schizosaccharomycetes.
NCBI_TaxID=4896;
[1]
SEQUENCE FROM N.A.
RA STRAIN=972;
RA MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgeson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeil C.,
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moser D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Furnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Foraburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrett B.G., Nurse P.,
RA "The genome sequence of Schizosaccharomycetes pombe."
RA Nature 415:871-880(2002).
CC -i- FUNCTION: Removes the amino-terminal methionine from nascent
CC proteins (By similarity).
CC -i- CATALYTIC ACTIVITY: Release of N-terminal amino acids,
CC preferentially methionine, from peptides and arylamides.
CC -i- COFACTOR: Cobalt; binds 2 ions per subunit (By similarity).
CC -i- SIMILARITY: Belongs to peptidase family M24.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AL023534; CAA19013.1; -
CC PIR: T40384; T40384.
CC HSP: P07906; 1C24.
CC MEROPS: M24.001; -.
CC GenDB SPombe; SPBC3E7.10; -.
CC InterPro: IPR001714; Pept_M24_MAP.
CC InterPro: IPR002467; Pept_M24A_MAP1.
CC InterPro: IPR000994; Peptidase M24.
CC Pfam: PF00557; Peptidase M24; 1.
CC PRINTS: PR00599; MAPEPTIDASE.
CC TIGRfam: TIGR00500; met_pdase_1; 1.
CC PROSITE: PS00680; MAP 1; FALSE_NEG.
KW Hydrolase; Aminopeptidase; Cobalt.
FT METAL 209 209 COBALT 2 (By similarity).
FT METAL 220 220 COBALT 1 AND 2 (By similarity).
FT METAL 289 289 COBALT 1 (By similarity).
FT METAL 322 322 COBALT 1 (By similarity).
FT METAL 353 353 COBALT 1 AND 2 (By similarity).
SQ SEQUENCE 379 AA; 42139 MW; D7A30E2886E3B248 CRC64;
Query Match 52.2%; Score 36; DB 1; Length 379;
Best Local Similarity 60.0%; Pred. No. 80;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 4 POTTNRDLDA 13
||||: ||

```

Db 148 PGTUDELDS 157

RESULT 23

CBP2_HUMAN STANDARD; PRT; 417 AA.

AC P48052; Q96A12; Q96QN3; --

DT 01-FEB-1996 (Rel. 33, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Carboxypeptidase A2 precursor (EC 3.4.17.15).

GN CPA2.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A., AND 3D-STRUCTURE MODELING.

RC TISSUE=Pancreas;

RX MEDLINE=95204457; PubMed=7896805;

RA Catasus L., Vendrell J., Aviles F.X., Carreira S., Puigserver A.,

RA Billeter M.;

RT "The sequence and conformation of human pancreatic

RT procarboxypeptidase A2. cDNA cloning, sequence analysis, and

RT three-dimensional model.";

RL J. Biol. Chem. 270:6651-6657(1995).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Brain, and Pancreas;

RX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,

RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk R.A.,

RA Villalón D.K., Munz D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,

RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length

RT human and mouse cDNA sequences";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RN [3]

RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).

RX MEDLINE=98111000; PubMed=9450539;

RA Reverter D., Garcia-Saez I., Catasus L., Vendrell J., Coll M.,

RA Aviles F.X.;

RT "Characterisation and preliminary X-ray diffraction analysis of human

RT pancreatic procarboxypeptidase A2.";

RL FEBS Lett. 420:7-10(1997).

RN [4]

RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).

RX MEDLINE=98046021; PubMed=9384570;

RA Garcia-Saez I., Reverter D., Vendrell J., Aviles F.X., Coll M.;

RT "The three-dimensional structure of human procarboxypeptidase A2.

RT Deciphering the basis of the inhibition, activation and intrinsic

RT activity of the zymogen.";

RL EMBO J. 16:6908-6913(1997).

CC -!- CATALYTIC ACTIVITY: Similar to that of carboxypeptidase A

CC (EC 3.4.17.1), but with a preference for bulkier C-terminal

CC residues.

CC -!- SIMILARITY: Belongs to peptidase family M14.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>

CC or send an email to license@isb-sib.ch).

CC

DR EMBL; U19977; AAA74425.1; --

DR EMBL; BC007009; AAH07009.1; --

DR EMBL; BC014571; AAH14571.1; --

DR EMBL; BC015140; AAH15140.1; --

DR PIR; A56171; A56171.

DR PDB; 1AYE; 13-JAN-99.

DR PDB; 1DT0; 12-JUL-00.

DR PDB; 1O6X; 30-JAN-03.

DR MEROPS; M14.002; --

DR GeneW; HGNC:2297; CPA2.

DR MIM; 600688; --

DR GO; GO:0004182; F:carboxypeptidase A activity; TAS.

DR GO; GO:0007039; P:vacuolar protein catabolism; TAS.

DR InterPro; IPR003146; Pept_M14A_propep.

DR InterPro; IPR000834; Peptidase_M14.

DR InterPro; IPR009020; Protease_inhib.

DR Pfam; PF02244; Propep_M14; 1.

DR PRINTS; PR00765; CRBOXYPTASEA.

DR SMART; SM00631; Zn_pept; 1.

DR PROSITE; PS00132; CARBOXYPEPT_ZN_1; 1.

DR PROSITE; PS00133; CARBOXYPEPT_ZN_2; 1.

KW Hydrolase; Carboxypeptidase; Metalloprotease; Zinc; Zymogen; Signal;

KW 3D-structure.

FT SIGNAL 1 16 POTENTIAL.

FT PROPEP 17 112 ACTIVATION PEPTIDE.

FT CHAIN 113 417 CARBOXYPEPTIDASE A2.

FT METAL 177 177 ZINC.

FT METAL 180 180 ZINC.

FT METAL 304 304 ZINC.

FT ACT SITE 378 378 NUCLEOPHILE.

FT DISULFID 246 269

FT DISULFID 318 352

FT CONFLICT 80 80

FT CONFLICT 302 302

SQ SEQUENCE 417 AA; 46828 MW; 3D9C073A98700A83 CRC64;

Query Match 52.2%; Score 36; DB 1; Length 417;

Best Local Similarity 75.0%; Pred. No. 89;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 GGPQTTSN 9

DB 260 GGFQASSN 267

RESULT 24

FIBA_HUMAN STANDARD; PRT; 866 AA.

AC P02671; Q9EX62; Q9UCH2;

DT 21-JUL-1986 (Rel. 01, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Fibrinogen alpha/alpha-E chain precursor [Contains: Fibrinopeptide A].

GN FGA.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A. (ISOFORM ALPHA-E).

RX MEDLINE=93090725; PubMed=1457396;

RA Fu Y., Weissbach L., Plant P.W., Oddoux C., Cao Y., Liang T.J.,

RA Roy S.N., Redman C.M., Grieninger G.;

RT "Carboxy-terminal-extended variant of the human fibrinogen alpha

RT subunit: a novel exon conferring marked homology to beta and gamma

RT subunits.";

- RL Biochemistry 31:11968-11972(1992).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM ALPHA-E).
 RA Chung D.W., Grieninger G.;
 RT "Fibrinogen DNA and protein sequences.";
 RL (In) Ebert R.F. (eds.),
 RL Index of variant human fibrinogens, pp.13-24, CRC Press,
 RL Boca Raton (1994).
 RN [3]
 RP SEQUENCE FROM N.A. (ALPHA-E; ALPHA), AND VARIANTS VAL-6; ALA-331 AND
 RP ALA-456.
 RA Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,
 RA Nickerson D.A.;
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 1-655 FROM N.A. (ISOFORM ALPHA-E).
 RC TISSUE=Liver;
 RX MEDLINE=91344740; PubMed=2102623;
 RA Chung D.W., Harris J.E., Davie E.W.;
 RT "Nucleotide sequences of the three genes coding for human
 RT fibrinogen.";
 RL Adv. Exp. Med. Biol. 281:39-48(1990).
 RN [5]
 RP SEQUENCE FROM N.A. (ISOFORM ALPHA).
 RX MEDLINE=83247396; PubMed=6575389;
 RA Kant J.A., Lord S.T., Crabtree G.R.;
 RT "Partial mRNA sequences for human A alpha, B beta, and gamma
 RT fibrinogen chains: evolutionary and functional implications.";
 RL Proc. Natl. Acad. Sci. U.S.A. 80:3953-3957(1983).
 RN [6]
 RP SEQUENCE OF 1-629 FROM N.A.
 RX MEDLINE=83283432; PubMed=6688355;
 RA Rixon M.W., Chan W.-Y., Davie E.W., Chung D.W.;
 RT "Characterization of a complementary deoxyribonucleic acid coding for
 RT the alpha chain of human fibrinogen.";
 RL Biochemistry 22:3237-3244(1983).
 RN [7]
 RP SEQUENCE OF 20-629.
 RA Henschen A., Lottepeich F., Southan C., Topfer-Petersen E.;
 RT "Human fibrinogen: sequence, sulfur bridges, glycosylation and some
 RT structural variants.";
 RL (In) Peeters H. (eds.);
 RL Protides of the biological fluids, Proc. 28th colloquium, pp.51-56,
 RL Pergamon Press, Oxford (1980).
 RN [8]
 RP SEQUENCE OF 20-629, AND DISULFIDE BONDS.
 RX MEDLINE=80088231; PubMed=518846;
 RA Watt K.W.K., Cottrell B.A., Strong D.D., Doolittle R.F.;
 RT "Amino acid sequence studies on the alpha chain of human fibrinogen.
 RT Overlapping sequences providing the complete sequence.";
 RL Biochemistry 18:5410-5416(1979).
 RN [9]
 RP SEQUENCE OF 110-156 FROM N.A.
 RX MEDLINE=84069777; PubMed=6689067;
 RA Imam A.M., Eaton M.A., Williamson R., Humphries S.;
 RT "Isolation and characterisation of cDNA clones for the A alpha- and
 RT gamma-chains of human fibrinogen.";
 RL Nucleic Acids Res. 11:7427-7434(1983).
 RN [10]
 RP SEQUENCE OF 605-644 FROM N.A. (ISOFORM ALPHA).
 RX MEDLINE=83254384; PubMed=6575700;
 RA Chung D.W., Rixon M.W., Que B.G., Davie E.W.;
 RT "Cloning of fibrinogen genes and their cDNA.";
 RL Ann. N.Y. Acad. Sci. 408:449-456(1983).
 RN [11]
 RP SEQUENCE OF 20-35.
 RA Blomback B., Blomback M., Grondahl N.J., Guthrie C., Hinton M.;
 RT "Studies on fibrinopeptides from primates.";
 RL Acta Chem. Scand. 19:1788-1789(1965).
 RN [12]
 RP CROSS-LINKING ACCEPTOR SITES.
 RX MEDLINE=80088230; PubMed=518845;
 RA Cottrell B.A., Strong D.D., Watt K.W.K., Doolittle R.F.;
 RT "Amino acid sequence studies on the alpha chain of human fibrinogen.
 RT Exact location of cross-linking acceptor sites.";
 RL Biochemistry 18:5405-5410(1979).
 RN [13]
 RP CROSS-LINKING ACCEPTOR SITES.
 RX MEDLINE=78130085; PubMed=632262;
 RA Fretto L.J., Ferguson E.W., Steinman H.M., McKee P.A.;
 RT "Localization of the alpha-chain cross-link acceptor sites of human
 RT fibrin.";
 RL J. Biol. Chem. 253:2184-2195(1978).
 RN [14]
 RP VARIANT, AND DISULFIDE BONDS.
 RX MEDLINE=76225080; PubMed=936108;
 RA Blomback B., Hessel B., Hogg D.;
 RT "Disulfide bridges in NH2-terminal part of human fibrinogen.";
 RL Thromb. Res. 8:639-658(1976).
 RN [15]
 RP REVIEW, EM STRUCTURE, POLYMERIZATION, AND LIGANDS.
 RX MEDLINE=84305751; PubMed=6383194;
 RA Doolittle R.F.;
 RT "Fibrinogen and fibrin.";
 RL Annu. Rev. Biochem. 53:195-229(1984).
 RN [16]
 RP CROSS-LINKING SITE FOR ALPHA-2-PLASMIN INHIBITOR.
 RX MEDLINE=87057190; PubMed=2877981;
 RA Kimura S., Aoki N.;
 RT "Cross-linking site in fibrinogen for alpha 2-plasmin inhibitor.";
 RL J. Biol. Chem. 261:15591-15595(1986).
 RN [17]
 RP PHOSPHORYLATION.
 RX MEDLINE=84104274; PubMed=6318767;
 RA Itarte E., Plana M., Guasch M.D., Martos C.;
 RT "Phosphorylation of fibrinogen by casein kinase 1.";
 RL Biochem. Biophys. Res. Commun. 117:631-636(1983).
 RN [18]
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 26-39.
 RX MEDLINE=92218459; PubMed=1560020;
 RA Martin P.D., Robertson W., Turk D., Huber R., Bode W., Edwards B.F.P.;
 RT "The structure of residues 7-16 of the A alpha-chain of human
 RT fibrinogen bound to bovine thrombin at 2.3-A resolution.";
 RL J. Biol. Chem. 267:7911-7920(1992).
 RN [19]
 RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS) OF 130-216.
 RX MEDLINE=97472408; PubMed=9333233;
 RA Spraggon G., Everse S.J., Doolittle R.F.;
 RT "Crystal structures of fragment D from human fibrinogen and its
 RT crosslinked counterpart from fibrin.";
 RL Nature 389:455-462(1997).
 RN [20]
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 130-216.
 RX MEDLINE=98292395; PubMed=9628725;
 RA Everse S.J., Spraggon G., Veerapandian L., Riley M., Doolittle R.F.;
 RT "Crystal structure of fragment double-D from human fibrin with two
 RT different bound ligands.";
 RL Biochemistry 37:8637-8642(1998).
 RN [21]
 RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 670-866.
 RX MEDLINE=98356117; PubMed=9689040;
 RA Spraggon G., Applegate D., Everse S.J., Zhang J.Z., Veerapandian L.,
 RA Redman C., Doolittle R.F., Grieninger G.;
 RT "Crystal structure of a recombinant alphaEC domain from human
 RT fibrinogen-420.";
 RL Proc. Natl. Acad. Sci. U.S.A. 95:9099-9104(1998).
 RN [22]
 RP X-RAY CRYSTALLOGRAPHY.
 RX MEDLINE=99175089; PubMed=10074346;
 RA Everse S.J., Spraggon G., Veerapandian L., Doolittle R.F.;
 RT "Conformational changes in fragments D and double-D from human
 RT fibrin(ogen) upon binding the peptide ligand Gly-His-Arg-Pro-amide.";
 RL Biochemistry 38:2941-2946(1999).
 RN [23]
 RP VARIANT KYOTO-2.
 RX MEDLINE=91300048; PubMed=2070049;

RA Yoshida N., Okuma M., Hirata H., Matsuda M., Yamazumi K., Asakura S.;
RT "Fibrinogen Kyoto II, a new congenitally abnormal molecule,
RT characterized by the replacement of A alpha proline-18 by leucine.";
RL Blood 78:149-153(1991).
RN [24]
RP VARIANT LIMA.
RX MEDLINE=92340680; PubMed=1634621;
RA Maekawa H., Yamazumi K., Muramatsu S., Kaneko M., Hirata H.,
RA Takahashi N., Arocha-Pinango C.L., Rodriguez S., Nagy H.,
RA Perez-Requejo J.L., Matsuda M.;
RT "Fibrinogen Lima: a homozygous dysfibrinogen with an A
RT alpha-arginine-141 to serine substitution associated with extra
RT N-glycosylation at A alpha-asparagine-139. Impaired fibrin gel
RT formation but normal fibrin-facilitated plasminogen activation
RT catalyzed by tissue-type plasminogen activator.";
RL J. Clin. Invest. 90:67-76(1992).
RN [25]
RP VARIANT CARACAS-2.
RX MEDLINE=91268018; PubMed=1675636;
RA Maekawa H., Yamazumi K., Muramatsu S., Kaneko M., Hirata H.,
RA Takahashi N., de Bosch N.B., Carvajal Z., Ojeda A.,
RA Arocha-Pinango C.L., Matsuda M.;
RT "An A alpha Ser-434 to N-glycosylated Asn substitution in a
RT dysfibrinogen, fibrinogen Caracas II, characterized by impaired
RT fibrin gel formation.";
RL J. Biol. Chem. 266:11575-11581(1991).
RN [26]
RP VARIANT DUSART.
RX MEDLINE=93232289; PubMed=8473507;
RA Koopman J., Haverkate F., Grimbergen J., Lord S.T., Moseson M.W.,
RA Diorio J.P., Siebenlist K.S., Legrand C., Soria J., Soria C.,
RA Caen J.P.;
RT "Molecular basis for fibrinogen Dusart (A alpha 554 Arg-->Cys) and
RT its association with abnormal fibrin polymerization and
RT thrombophilia.";
Query Match 52.2%; Score 36; DB 1; Length 866;
Best Local Similarity 75.0%; Pred. No. 2e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 3 GPGTTSNR 10
DB 301 GPGTGNR 308
RESULT 25
FTSK RHIME STANDARD; PRT; 881 AA.
ID FTSK RHIME
AC Q92L89;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE DNA translocase ftsK.
GN FTSK OR R03191 OR SMC03808.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021,
RX MEDLINE=21396507; PubMed=11481430;
RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Gohl T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
RA Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,
RA Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.;
RT "Analysis of the chromosome sequence of the legume symbiont
RT Sinorhizobium meliloti strain 1021.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
CC -1- FUNCTION: DNA motor protein, which is both required to move DNA
CC out of the region of the septum during cell division and for the
CC septum formation. Tracks DNA in an ATP-dependent manner by

CC generating positive supercoils in front of it and negative
CC supercoils behind it (By similarity).
CC -1- SUBUNIT: Homohexamer. This suggests the formation of a ring
CC similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Located at the
CC septum. The large C-terminal part of the protein is cytoplasmic
CC (Potential).
CC -1- SIMILARITY: Contains 1 FtsK domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL591793; CAC47770.1; -;
DR HAMAP; MF_01809; -; 1;
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR002543; FtsK_SpoIIIE.
DR Pfam; PF01580; FtsK_SpoIIIE; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS50901; FTSK; 1.
KW Chromosome partition; Cell division; ATP-binding; DNA-binding;
KW Transmembrane; Complete proteome.
FT TRANSMEM 21 43 POTENTIAL.
FT TRANSMEM 78 100 POTENTIAL.
FT TRANSMEM 113 135 POTENTIAL.
FT DOMAIN 514 733 FTSK.
FT NP_BIND 531 538 ATP (POTENTIAL).
SQ SEQUENCE 881 AA; 95549 MW; 28E890501D72565F CRC64;
Query Match 52.2%; Score 36; DB 1; Length 881;
Best Local Similarity 58.3%; Pred. No. 2e+02;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 2 GPGTTSNRDLA 13
DB 802 GPGAGTNSLADS 813
Search completed: September 24, 2004, 01:55:27
Job time : 14.7925 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 01:47:17 ; Search time 62.3019 Seconds
(without alignments)
65.836 Million cell updates/sec

Title: US-10-088-639a-2_COPY_226_238

Perfect score: 69

Sequence: 1 EGGPGTTSRLDA 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database :

SPTREMBL 25: *

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phase:*

10: sp_plant:*

11: sp_rodent:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

15: sp_rvirus:*

16: sp_bacteriap:*

17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	63.8	419	16	Q7UNR6 rhodopirell
2	44	63.8	3926	4	Q9UPA5 homo sapien
3	43	62.3	749	5	Q9Y149 drosophila
4	42	60.9	313	16	Q89GJ0 bradyrhizob
5	42	60.9	503	11	Q8BXM8 mus musculus
6	42	60.9	541	11	Q8QZT7 mus musculus
7	42	60.9	579	11	Q8C0G3 mus musculus
8	42	60.9	588	11	Q8CH42 mus musculus
9	42	60.9	598	11	Q8C078 mus musculus
10	42	60.9	1513	13	Q90Z70 brachydanio
11	40	58.0	105	5	Q8WPUE trypanosoma
12	40	58.0	249	16	Q84893 salmonella
13	40	58.0	283	6	Q9XSZ8 cercopithec
14	40	58.0	313	16	Q9RL33 streptomyce
15	40	58.0	351	16	Q9XAB8 streptomyce
16	40	58.0	757	16	Q82WD2 nitrosomona

17	39	56.5	67	16	Q92MH7	Q92mh7 rhizobium m
18	39	56.5	130	16	Q883M3	Q883m3 pseudomonas
19	39	56.5	146	10	Q8LR30	Q8lr30 oryza sativ
20	39	56.5	193	15	Q8Q2G1	Q8q2g1 human immun
21	39	56.5	193	16	Q7VRN6	Q7vrn6 candidatus
22	39	56.5	278	11	Q8CG89	Q8cg89 mus musculus
23	39	56.5	300	6	Q9GLY9	Q9gly9 macaca fasc
24	39	56.5	301	16	Q98MFS	Q98mf5 rhizobium l
25	39	56.5	302	5	Q9VVF5	Q9vvf5 drosophila
26	39	56.5	385	16	Q81ZWH	Q81zwh streptomyce
27	39	56.5	396	4	Q86K21	Q86k21 homo sapien
28	39	56.5	398	11	Q7TQJ8	Q7tcj8 mus musculus
29	39	56.5	422	5	Q9SSE1	Q9sse1 drosophila
30	39	56.5	454	16	Q834L7	Q834l7 enterococcu
31	39	56.5	473	10	Q84WQ1	Q84wq1 arabidopsis
32	39	56.5	499	4	Q86WC2	Q86wc2 homo sapien
33	39	56.5	615	16	Q9KYX0	Q9kyx0 streptomyce
34	39	56.5	750	11	Q8BVV4	Q8bv4 mus musculus
35	39	56.5	770	16	Q81Z23	Q81zz3 nitrosomona
36	39	56.5	775	11	Q7TQJ9	Q7tcj9 mus musculus
37	39	56.5	776	2	Q9L8K6	Q9l8k6 mycobacteri
38	39	56.5	932	16	Q92N47	Q92n47 rhizobium m
39	39	56.5	1112	16	Q92UV3	Q92uv3 rhizobium m
40	39	56.5	1121	5	Q9UAN9	Q9uan9 caenorhabdi
41	39	56.5	1123	5	Q8ITW0	Q8itw0 caenorhabdi
42	39	56.5	1157	5	Q9VP22	Q9vp22 drosophila
43	39	56.5	1157	5	Q8T9E1	Q8t9e1 drosophila
44	39	56.5	1253	5	Q8MSP0	Q8msp0 drosophila
45	39	56.5	1253	5	Q9VPP2	Q9vpp2 drosophila
46	39	56.5	1317	11	Q7TQA1	Q7tqa1 mus musculus
47	39	56.5	1317	11	Q88778	Q88778 rattus norv
48	39	56.5	3938	11	Q88G54	Q88g54 rhizobium l
49	38	55.1	109	6	Q9BGZ9	Q9bgz9 macaca fasc
50	38	55.1	201	15	Q9E4Q4	Q9e4q4 human immun
51	38	55.1	201	15	Q9E4Q9	Q9e4q9 human immun
52	38	55.1	201	15	Q9E4Q7	Q9e4q7 human immun
53	38	55.1	201	15	Q9E4R0	Q9e4r0 human immun
54	38	55.1	201	15	Q9E4Q2	Q9e4q2 human immun
55	38	55.1	201	15	Q9E4Q5	Q9e4q5 human immun
56	38	55.1	201	15	Q9E4Q6	Q9e4q6 human immun
57	38	55.1	201	15	Q9E4R2	Q9e4r2 human immun
58	38	55.1	201	15	Q9E4R1	Q9e4r1 human immun
59	38	55.1	207	11	Q9JL32	Q9jl32 mus musculus
60	38	55.1	227	5	Q9N8S8	Q9n8s8 trypanosoma
61	38	55.1	227	10	Q8LAJ4	Q8laj4 arabidopsis
62	38	55.1	243	16	Q88V99	Q88v99 lactobacill
63	38	55.1	249	11	Q8R058	Q8r058 mus musculus
64	38	55.1	257	16	Q7UHT9	Q7uht9 rhodopirell
65	38	55.1	263	16	Q89KQ5	Q89kq5 bradyrhizob
66	38	55.1	275	10	Q9FJM1	Q9fjm1 arabidopsis
67	38	55.1	312	16	Q8UC11	Q8uc11 agrobacteri
68	38	55.1	337	16	Q983J2	Q983j2 rhizobium l
69	38	55.1	409	16	Q828M2	Q828m2 streptomyce
70	38	55.1	516	16	Q7TYP2	Q7typ2 mycobacteri
71	38	55.1	623	16	Q88KI6	Q88ki6 pseudomonas
72	38	55.1	627	16	Q87Y50	Q87ys0 pseudomonas
73	38	55.1	667	16	P71749	P71749 mycobacteri
74	38	55.1	749	3	Q96UA7	Q96ua7 neurospora
75	38	55.1	1219	3	Q53785	Q53785 streptomyce
76	38	55.1	1872	5	Q9XTE2	Q9xte2 caenorhabdi
77	37.5	54.3	505	1	Q50530	Q50530 methanobact
78	37	53.6	52	10	Q9ZSF9	Q9zsf9 lycium ande
79	37	53.6	94	17	Q8PRS1	Q8prs1 methanosarc
80	37	53.6	139	16	Q9RTA8	Q9rta8 deinococcus
81	37	53.6	159	10	Q941P5	Q941p5 malus domes
82	37	53.6	204	16	Q88D47	Q88d47 pseudomonas
83	37	53.6	229	3	Q43008	Q43008 schizosacch
84	37	53.6	251	9	Q7Y2C7	Q7y2c7 bacterioph
85	37	53.6	256	16	Q89JU5	Q89ju5 bradyrhizob
86	37	53.6	283	4	Q81Y45	Q81y45 homo sapien
87	37	53.6	287	16	Q7UMG2	Q7umg2 rhodopirell
88	37	53.6	290	2	O68115	O68115 rhodobacter
89	37	53.6	290	16	Q92N03	Q92n03 rhizobium m

```

90 37 53.6 294 4 Q86X98
91 37 53.6 294 11 Q9B8U4
92 37 53.6 297 16 Q89T4
93 37 53.6 308 16 Q8NU61
94 37 53.6 325 16 Q8FUC8
95 37 53.6 326 16 Q7V067
96 37 53.6 329 16 Q88DP8
97 37 53.6 352 6 Q95LJ2
98 37 53.6 353 10 Q40366
99 37 53.6 360 5 Q19879
100 37 53.6 377 16 Q8BMO5
101 37 53.6 385 16 Q8CH82
102 37 53.6 391 10 Q8EX70
103 37 53.6 398 5 Q8IR93
104 37 53.6 402 5 Q9FJ89
105 37 53.6 433 5 Q9VZ20
106 37 53.6 538 8 Q8W9R3
107 37 53.6 543 4 Q9H3E4
108 37 53.6 586 12 Q9W8B7
109 37 53.6 604 2 Q00509
110 37 53.6 616 16 Q7UJ59
111 37 53.6 616 17 Q9YBX7
112 37 53.6 626 11 Q921U9
113 37 53.6 691 4 Q7Z322
114 37 53.6 1119 4 Q75762
115 37 53.6 1218 2 Q9FB24
116 37 53.6 1756 16 Q89I18
117 37 53.6 1777 4 Q9ULH0
118 37 53.6 2203 5 Q963L8
119 37 53.6 2618 5 Q9VP88
120 37 53.6 2796 11 Q8OT79
121 37 53.6 3357 12 Q8ODQ3
122 37 53.6 3357 12 Q8ODQ2
123 37 53.6 3501 4 Q7Z407
124 36 52.2 164 10 Q7XUH9
125 36 52.2 178 16 Q9AN15
126 36 52.2 192 16 Q87V70
127 36 52.2 204 15 Q9E417
128 36 52.2 216 3 Q9PB19
129 36 52.2 221 13 Q7Z262
130 36 52.2 221 16 Q8PIE9
131 36 52.2 222 16 Q832H4
132 36 52.2 227 16 Q7TUX5
133 36 52.2 234 16 Q9AK73
134 36 52.2 240 16 Q9PFM2
135 36 52.2 250 5 Q812H9
136 36 52.2 256 4 Q96MK8
137 36 52.2 288 16 Q89CV0
138 36 52.2 290 2 Q30832
139 36 52.2 294 16 Q89OD8
140 36 52.2 313 5 Q61525
141 36 52.2 350 10 Q93XK6
142 36 52.2 351 10 Q24081
143 36 52.2 356 16 Q8Y230
144 36 52.2 363 15 Q7ZN27
145 36 52.2 379 16 Q7U8C9
146 36 52.2 394 2 Q9ZAM0
147 36 52.2 399 5 Q9NG02
148 36 52.2 400 16 Q92TL0
149 36 52.2 401 2 Q845J8
150 36 52.2 402 2 Q84H55

```

ALIGNMENTS

```

RESULT 1
Q7UNR6 PRELIMINARY; PRT; 419 AA.
ID Q7UNR6;
AC Q7UNR6;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

```

```

DE Aminotransferase Nifs-putative Nifs homolog (EC 4.4.1.-).
GN NIFS OR RB7424.
OS Rhodopirellula baltica.
OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Pirellula.
OX NCBI_TaxID=117;
RN [1]_TaxID=117;
RP SEQUENCE FROM N.A.
RC STRAIN=1;
RX MEDLINE=22735913; PubMed=12835416;
RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,
RA Schlesner H., Amann R., Reinhardt R.;
RT "Complete genome sequence of the marine planctomycete Pirellula sp.
RT strain 1."
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
RW EMBL; BX294146; CAD75352.1; -.
KW Aminotransferase; Lyase; Transferase; Complete proteome.
SQ SEQUENCE 419 AA; 43485 MW; B91E67B8F2C62D79 CRC64;

Query Match 63.8%; Score 44; DB 16; Length 419;
Best Local Similarity 66.7%; Pred. No. 22;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 GGGTTSNRDLA 13
DB 392 GGGTTSNRDLA 403

RESULT 2
Q9UPA5 PRELIMINARY; PRT; 3926 AA.
ID Q9UPA5;
AC Q9UPA5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Neuronal double zinc finger protein.
GN ZNF231.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP SEQUENCE FROM N.A.
RX MEDLINE=99026225; PubMed=9806829;
RA Hashida H., Goto J., Zhao N., Takahashi N., Hirai M., Kanazawa I.,
RA Sakaki Y.;
RT "Cloning and mapping of ZNF231, a novel brain-specific gene encoding
RT neuronal double zinc finger protein whose expression is enhanced in a
RT neurodegenerative disorder, multiple system atrophy."
RL Genomics 54:50-58(1998).
DR EMBL; AF052224; AAC83555.1; -.
DR Genew; HGNC:11117; BSN.
DR GO; GO:0005634; C:nucleus; TAS.
DR GO; GO:0007268; P:synaptic transmission; TAS.
DR InterPro; IPR008899; Znf_piccolo.
DR Fram; PF05715; Zf_piccolo; 2.
SQ SEQUENCE 3926 AA; 416458 MW; 75CDS6EADD437343 CRC64;

Query Match 63.8%; Score 44; DB 4; Length 3926;
Best Local Similarity 80.0%; Pred. No. 2.3e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 GFGTTSNRDL 12
DB 72 GFGTTSNRDL 81

RESULT 3
Q9Y149 PRELIMINARY; PRT; 749 AA.
ID Q9Y149;
AC Q9Y149;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)

```

01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 BCDNA:GH03922 protein.
 ARC105 OR BCDNA:GH03922 OR CG4184.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RX MEDLINE=2019606; PubMed=10731132;
 RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Ananides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandall M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champagne M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolehakov S.,
 RA Borkova D., Borchan M.R., Bouck J., Brokstein P., Brotier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong P.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster";
 RL Science 287:2185-2195(2000).
 RN
 RP SEQUENCE FROM N.A.
 RA Rubin G.M., Wan K.H., Harvey D., Lewis S.E., Brokstein P., Tsang G.,
 RA Agbayani A., Arcata T.T., Baxter E., Blazej R.G., Butenhoff C.,
 RA Champe M., Chavez C., Chew M., Doyle C.M., Farfan D.E., Frise E.,
 RA Galle R., George R.A., Harris N.L., Hoskins R.A., Evans-Holm M.,
 RA Houston K.A., Hummasti S.R., Kim E., Li P., Moshrefi M., Pacleb J.M.,
 RA Park S., Sequeira A., Sethi H., Snir E., Svirskaas R.R., Weinburg T.,
 RA Celnik S.E.;
 RT "Full length Drosophila melanogaster cDNA sequence";
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE003589; AAF51490.1; -;
 DR ENBL; AFI45620; AAD38595.1; -;
 DR FlyBase; FBgn0027592; Arc105.
 SQ SEQUENCE 749 AA; 80526 MW; 42375AEB7C13A1DF CRC64;

Query Match 62.3%; Score 43; DB 5; Length 749;
 Best Local Similarity 66.7%; Pred. No. 62;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 GGGTTSNRDLA 13
 ||||| | | |

Db 127 GGGTASNLQ 138
 RESULT 4
 Q89GJ0 PRELIMINARY; PRT; 313 AA.
 ID Q89GJ0;
 AC Q89GJ0;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE ABC transporter permease protein.
 GN BUL6355.
 OS Bradyrhizobium japonicum.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Bradyrhizobiaceae; Bradyrhizobium.
 NCBI_TaxID=375;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=USDA 110;
 RX MEDLINE=22484998; PubMed=12597275;
 RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
 RA Sasanoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
 RA Kohara M., Matsumoto M., Shimo S., Tsuruoka H., Wada T., Yamada M.,
 RA Tabata S.;
 RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
 Bradyrhizobium japonicum USDA110";
 RL DNA Res. 9:189-197(2002).
 DR EMBL; AP005958; BAC51620.1; -;
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005215; F:transporter activity; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR000515; BPD transp.
 DR Pfam; PF00528; BPD transp. 1.
 DR PROSITE; PS00402; BPD_TRANS_INN_MEMBER; 1.
 KW Complete proteome.
 SQ SEQUENCE 313 AA; 34112 MW; AE97D4548E8394A0 CRC64;

Query Match 60.9%; Score 42; DB 16; Length 313;
 Best Local Similarity 58.3%; Pred. No. 38;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EGGPGTTSNRDL 12
 :|||: | | |

Db 259 QGGPGNSNTLN 270

RESULT 5
 Q8BXN8 PRELIMINARY; PRT; 503 AA.
 ID Q8BXN8;
 AC Q8BXN8;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DE CA/calmodulin-dependent protein kinase beta.
 GN CAMK2 OR G330570N16R1K.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Retina;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The PANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 60,770 full-length cDNAs";
 RL Nature 420:563-573(2002).
 DR EMBL; AK044660; BAC32023.1; -;
 DR MGD; MGI:2444812; Camk2.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004674; F:protein-tyrosine kinase activity; IEA.
 DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.

DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR002290; Ser_thr_pkinase.
 DR InterPro; IPR008271; Ser_thr_pkin_AS.
 DR InterPro; IPR001245; Tyr_pkinase.
 DR pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00220; S_TKC; 1.
 DR SMART; SM00219; TyrKC; 1.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.
 DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
 SQ SEQUENCE 503 AA; 55501 MW; 8C596DEB39649F42 CRC64;
 Query Match 60.9%; Score 42; DB 11; Length 503;
 Best Local Similarity 58.3%; Pred. No. 62;
 Matches 7; Conservative. 3; Mismatches 2; Indels 0; Gaps 0;
 QY 1 EGGPGTTSNRLD 12
 Db 105 QGGPASSNSLD 116
 RESULT 6
 Q8QZT7 PRELIMINARY; PRT; 541 AA.
 AC Q8QZT7;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Similar to Ca/calmodulin-dependent protein kinase beta (CaM-
 DE kinase kinase beta).
 DE CAMKK2 OR 6330570N16RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 CC 1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL; BC023103; AAH23103.1; -.
 DR MGD; MGI:2444812; Camkk2.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004674; P:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0016740; P:transferase activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR002290; Ser_thr_pkinase.
 DR InterPro; IPR008271; Ser_thr_pkin_AS.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00220; S_TKC; 1.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.
 DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 541 AA; 59616 MW; FEB2385A4D793FF2 CRC64;
 Query Match 60.9%; Score 42; DB 11; Length 541;
 Best Local Similarity 58.3%; Pred. No. 66;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 1 EGGPGTTSNRLD 12
 Db 105 QGGPASSNSLD 116
 RESULT 7
 Q8COG3 PRELIMINARY; PRT; 579 AA.
 ID Q8COG3
 AC Q8COG3;

DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Ca/calmodulin-dependent protein kinase kinase beta.
 GN CAMKK2 OR 6330570N16RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=22354683; PubMed=12466651;
 RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573 (2002).
 DR EMBL; AK031399; BAC27387.1; -.
 DR MGD; MGI:2444812; Camkk2.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004674; P:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0004713; P:protein-tyrosine kinase activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR002290; Ser_thr_pkinase.
 DR InterPro; IPR008271; Ser_thr_pkin_AS.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00220; S_TKC; 1.
 DR SMART; SM00219; TyrKC; 1.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.
 DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
 SQ SEQUENCE 579 AA; 63477 MW; A79C6F3E1CE6DDCA CRC64;
 Query Match 60.9%; Score 42; DB 11; Length 579;
 Best Local Similarity 58.3%; Pred. No. 71;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 1 EGGPGTTSNRLD 12
 Db 105 QGGPASSNSLD 116
 RESULT 8
 Q8CH42 PRELIMINARY; PRT; 588 AA.
 ID Q8CH42;
 AC Q8CH42;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Ca/calmodulin-dependent protein kinase kinase beta.
 GN CAMKK2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Carra S.; Harvey M.; Tacedda F.; Barden N.;
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF453383; AAN75696.1; -.
 DR MGD; MGI:2444812; Camkk2.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004674; P:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0004713; P:protein-tyrosine kinase activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR002290; Ser_thr_pkinase.
 DR InterPro; IPR008271; Ser_thr_pkin_AS.
 DR InterPro; IPR001245; Tyr_pkinase.

DR Pfam; PF00069; kinase; 1.
 DR ProDom; PD00001; Prot kinase; 1.
 DR SMART; SM00220; S_TKc; 1.
 DR SMART; SM00219; TyrKc; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW Kinase.
 SQ SEQUENCE 588 AA; 64753 MW; 8026EE485C3BA06A CRC64;

Query Match 60.9%; Score 42; DB 11; Length 588;
 Best Local Similarity 58.3%; Pred. No. 72;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EGGPGTTSNRDL 12
 :||| :||| ||
 Db 105 QGGPASSNSLD 116

RESULT 9

ID Q8C078 PRELIMINARY; PRT; 588 AA.
 AC Q8C078;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE CA+calmodulin-dependent protein kinase beta.
 GN CAMK2 OR 6330570N16R1K.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Medulla oblongata;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs";
 RL Nature 420:563-573 (2002).
 DR EMBL; AK032070; BAC27681.1; -.
 DR MGD; MGI:2444812; Camk2.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR000719; Prot kinase.
 DR InterPro; IPR002290; Ser_thr_kinase.
 DR InterPro; IPR008271; Ser_thr_kin_AS.
 DR Pfam; PF00069; kinase; 1.
 DR ProDom; PD00001; Prot kinase; 1.
 DR SMART; SM00220; S_TKc; 1.
 DR SMART; SM00219; TyrKc; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 SQ SEQUENCE 588 AA; 64598 MW; 95B13E1C28FC614F CRC64;

Query Match 60.9%; Score 42; DB 11; Length 588;
 Best Local Similarity 58.3%; Pred. No. 72;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EGGPGTTSNRDL 12
 :||| :||| ||
 Db 105 QGGPASSNSLD 116

RESULT 10

ID Q90Z70 PRELIMINARY; PRT; 1513 AA.
 AC Q90Z70;

DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Roundabout2.
 GN ROBO2.
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21270059; PubMed=11376489;
 RA Lee J.S., Ray R., Chien C.B.;
 RT "Cloning and expression of three zebrafish roundabout homologs suggest
 RT roles in axon guidance and cell migration.";
 RL Dev. Dyn. 221:216-230 (2001).
 DR EMBL; AF337035; AAKS8427.1; -.
 DR ZFIN; ZDB-GENE-001019-1; robo2.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
 DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
 DR InterPro; IPR003961; FN_III.
 DR InterPro; IPR008957; FN_III-like.
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003598; Ig_c2.
 DR Pfam; PF00041; fn3; 3.
 DR Pfam; PF00047; ig; 5.
 DR SMART; SM00060; FN3; 3.
 DR SMART; SM00408; IGC2; 5.
 DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
 DR PROSITE; PS00835; IG_LIKE; 5.
 KW Immunoglobulin domain.
 SQ SEQUENCE 1513 AA; 165181 MW; D1743BACCC089F0 CRC64;

Query Match 60.9%; Score 42; DB 13; Length 1513;
 Best Local Similarity 72.7%; Pred. No. 1.9e+02;
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 GPGTTSNRDLA 13
 :||| :||| ||
 Db 1350 GPGTTPNRDA 1360

RESULT 11

ID Q8WPUG PRELIMINARY; PRT; 105 AA.
 AC Q8WPUG;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 GN 13J3.03.
 OS Trypanosoma brucei.
 OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
 OX NCBI_TaxID=5691;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=427;
 RA Berriman M., Hall N., Shearer K., Brinkard F., Tiwari B., Isobe T.,
 RA Bowman S., Cotton C., Clark L., Cross G.A.M., Hoek M., Zanders T.,
 RA Berberof M., Borst P., Rudenko G.;
 RT "The architecture of Variant Surface Glycoprotein gene expression
 RT sites in Trypanosoma brucei.";
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases. ;
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=427;
 RA Zeng C., Zhao B., Hierl M., Catanese J., Gerrard C., Melville S.E.,
 RA Hoek M., Navarro M., Cross G.A.M., El-Sayed N., Berberof M.,
 RA Rudenko G., Borst P., de Jong P.;
 RT "Construction and Characterization of Three Bacterial Artificial

```

RT Chromosome Libraries for Trypanosoma brucei."
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=427;
RX MEDLINE=21429110; PubMed=11543629;
RA Zeng C., Kouprina N., Zhu B., Cairo A., Hoek M., Cross G.A.M.,
RA Ooegawa K., Larionov V., de Jong P.;
RT "Large-insert BAC/YAC libraries for selective re-isolation of genomic
RT regions by homologous recombination in yeast.";
RL Genomics 77:27-34(2001).
DR EMBL; AL670322; CAD21450.1; -.
KW Hypothetical protein.
SQ SEQUENCE 105 AA; 11507 MW; 1C875F8030278B66 CRC64;

Query Match 58.0%; Score 40; DB 5; Length 105;
Best Local Similarity 66.7%; Pred. No. 28;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 EGGPGTTSNRDL 12
DB 77 EGGGKTKNRKD 88

RESULT 12
084893 PRELIMINARY; PRT; 249 AA.
AC 084893;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Putative DNA replication protein (Hypothetical protein) (Gifsy-2
DE prophage, ATPase involved in DNA replication initiation).
GN T1917 OR STM2625 OR STM1015 OR STY1023.
OS Salmonella typhi, and
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=601, 602;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhi; STRAIN=Ty2 / ATCC 700931;
RX MEDLINE=22531367; PubMed=12644504;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA Burland V., Kodoyanni V., Schwartz D.C., Blattner F.R.;
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
RT and CT18.";
RL J. Bacteriol. 185:2330-2337(2003).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhimurium; STRAIN=LT2;
RA Figueroa-Bossi N., Bossi L.;
RT "Prophage genes in Salmonella.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhimurium; STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.";
RL Nature 413:852-856(2001).
RN [4]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhi; STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,

```

```

RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrell B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18.";
RL Nature 413:848-852(2001).
DR EMBL; AE016840; AAC069533.1; -.
DR EMBL; AF001386; AAC26072.1; -.
DR EMBL; AE008819; AAL21519.1; -.
DR EMBL; AE008743; AAL19949.1; -.
DR EMBL; AL627268; CAD05417.1; -.
DR PIR; T03011; T03011.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR002198; ADH_Short.
DR InterPro; IPR002611; ISEB_ATPbind.
DR Pfam; PF01695; ISEB; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00061; ADH_SHORT; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 249 AA; 27568 MW; 6DSD99B79B696BA7 CRC64;

Query Match 58.0%; Score 40; DB 16; Length 249;
Best Local Similarity 66.7%; Pred. No. 67;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 GGPCTTSNRDLA 13
DB 109 GGPCTGKNHLAA 120

RESULT 13
09XSZ8 PRELIMINARY; PRT; 283 AA.
AC 09XSZ8;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Hveas.
DE HVEAS.
OS Cercopithecus aethiops (Green monkey) (Grivet).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheidae; Cercopithecus.
OX NCBI_TaxID=9534;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=99296730; PubMed=10366573;
RA Foster T.P., Chouljenko V.N., Kousoulas K.G.;
RT "Functional characterization of the HveA homolog specified by African
RT green monkey kidney cells with a herpes simplex virus expressing the
RT green fluorescence protein.";
RL Virology 258:365-374(1999).
DR EMBL; AF147720; AAD37381.1; -.
DR HSSP; Q92956; LJWA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004888; F:transmembrane receptor activity; IEA.
DR GO; GO:0006915; P:apoptosis; IEA.
DR GO; GO:0006955; P:immune response; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR008063; Fas_receptor.
DR InterPro; IPR001368; TNFR_c6.
DR Pfam; PF00020; TNFR_C6; 3.
DR PRINTS; PR01680; FASRECEPTOR.
DR SMART; SM00208; TNFR; 3.
DR PROSITE; PS00652; TNFR_NGFR_1; 1.
DR PROSITE; PS00500; TNFR_NGFR_2; 2.
SQ SEQUENCE 283 AA; 30199 MW; 397951C6617FE3AA CRC64;

```

Query Match 58.0%; Score 40; DB 6; Length 283;
Best Local Similarity 70.0%; Pred. No. 77;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EGGPGTTSNR 10
| | | | | | | |
DB 192 EAGPGTSSR 201

RESULT 14
Q9RL33 PRELIMINARY; PRT; 313 AA.
AC Q9RL33;
DT 01-WAY-2000 (TREMBLrel. 13, Created)
DT 01-WAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Binding-protein dependent transport protein.
GN SC00454 OR SCF51A.32.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Murphy L., Harris D.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kieser H.M., Denapaita D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
coelicolor A3(2).";
RL Nature 417:141-147(2002).
CC -!- FUNCTION: PART OF A BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM.
CC PROBABLY RESPONSIBLE FOR THE TRANSLLOCATION OF THE SUBSTRATE ACROSS
CC THE MEMBRANE (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: WITH INTEGRAL MEMBRANE COMPONENTS OF OTHER BINDING-
CC PROTEIN-DEPENDENT TRANSPORT SYSTEMS.
DR EMBL; AL939105; CAB56684.1; --
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000515; BPD_transp.
DR Pfam; PF00528; BPD_transp.1.
DR PROSITE; PS00402; BPD_TRANSF_INN_MEMBR; 1.
KW Transmembrane; Transport; Complete proteome.
SQ SEQUENCE 313 AA; 33365 MW; 2AE7FAEDF5AC56B0 CRC64;

Query Match 58.0%; Score 40; DB 16; Length 313;
Best Local Similarity 63.6%; Pred. No. 85;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EGGPGTTSNRL 11
| | | | | | | |
DB 260 QGGPGTSTTL 270

RESULT 15
Q9XAB8 PRELIMINARY; PRT; 351 AA.
AC Q9XAB8;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Hypothetical protein SC01938.
GN SC01938 OR SCC22.20.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Seeger K., Harris D.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Bentley S.D., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kieser H.M., Denapaita D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL939110; CAB50763.1; --
DR FIR; T36010; T36010.
DR InterPro; IPR004555; OpcA.
DR TIGRFAMs; TIGR00534; OpcA; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 351 AA; 37677 MW; EE133DFBF2651577 CRC64;

Query Match 58.0%; Score 40; DB 16; Length 351;
Best Local Similarity 63.6%; Pred. No. 96;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 GGGPGTTSNRLD 12
| | | | | | | |
DB 233 GGGPGTAVRMD 243

```

RESULT 16
Q82WD2 ID Q82WD2 PRELIMINARY; PRT; 757 AA.
AC Q82WD2;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DE 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE TonB-dependent receptor protein.
GN NE0758.
OS Nitrosomonas europaea.
OC Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;
OC Nitrosomonadaceae; Nitrosomonas.
OX NCBI_TaxID=915;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 19718 / IPO 14298;
RA Chain P., Lamerdin J.E., Larimer F.W., Regala W., Lao V., Land M.,
HAUSER L., Hooper A.B., Klotz M.G., Norton J., Sayavedra-Soto L.A.,
ARCIERO D.M., HOMMES N.G., WHITTAKER M.M., ARP D.J.;
RT "Complete genome sequence of the ammonia-oxidizing bacterium and
obligate chemolithoautotroph Nitrosomonas europaea.";
RL J. Bacteriol. 185:2759-2773(2003).
DR EMBL: BX321858; CAD84669.1; -
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0004872; F:receptor activity; IEA.
DR GO: GO:0005215; F:transporter activity; IEA.
DR GO: GO:0006810; P:transport; IEA.
DR InterPro: IPR000531; TonB_boxC.
DR Pfam: PF00593; TonB_dep_REC; 1.
DR PROSITE: PS01156; TONB_DEPENDENT_REC_2; 1.
DR Receptor; Complete proteome.
KW SEQUENCE 757 AA; 83756 MW; CD48F4FD5B2B891 CRC64;

Query Match 58.0%; Score 40; DB 16; Length 757;
Best Local Similarity 60.0%; Pred. No. 2; Mismatches 0; Indels 0; Gaps 0;
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGGPGTTSNR 10
:|||||:
DB 384 DGGPGTSAHR 393

RESULT 17
Q92MH7 ID Q92MH7 PRELIMINARY; PRT; 67 AA.
AC Q92MH7;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Hypothetical transmembrane protein SMC00732.
GN R02640 OR SMC00732.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=1021;
RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
BOISTARD P., BECKER A., BOUTRY M., CADIEU E., DREANO S., GLOUX S.,
GODRIE T., GOFFEAU A., KAHN D., KIAS E., LELAURE V., MASUY D.,
POHL T., PORTETTELLE P., PUEHLER A., PURNELLE B., RAMSPERGER U.,
RENARD C., THEBAULT D., VANDENBOL M., WEIDNER S., GALIBERT F.;
RT "Analysis of the chromosome sequence of the legume symbiont
Sinorhizobium meliloti strain 1021.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
DR EMBL: AL591791; CAC47219.1; -
KW Hypothetical protein; Complete proteome.
KW SEQUENCE 67 AA; 7030 MW; 781F9A0C593PBF85 CRC64;

Query Match 56.5%; Score 39; DB 16; Length 67;
Best Local Similarity 63.6%; Pred. No. 26; Mismatches 2; Indels 0; Gaps 0;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EGGPGTTSNRL 11
:|||||:
DB 29 KGGSGNTSNKL 39

RESULT 18
Q883M3 ID Q883M3 PRELIMINARY; PRT; 130 AA.
AC Q883M3;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Conserved hypothetical protein.
GN PSPTO2333.
OS Pseudomonas syringae (pv. tomato).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=323;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=DC3000;
RA Buell R., Joardar V., Khouri H., Fedorova N., Tran B., Russell D.,
BERRY K., UTTERBACK T., VAN AKEN S., FELDBLYUM T., GWINN M.,
DODSON R., DEBOY R., DURKIN A., KOLONAY J., MADUPU R., DAUGHERTY S.,
BRINKAC L., BEANAN M., HAFT D., SELENGUT J., NELSON W., DAVIDSEN T.,
WHITE O., FRASER C., COLLIER A.;
RT "Complete sequence of Pseudomonas syringae.";
RL EMBL: AF016864; AAO55845.1; -
DR TIGR; PSPTO2333; -
KW Hypothetical protein; Complete proteome.
KW SEQUENCE 130 AA; 14087 MW; 8B290C4A5E586010 CRC64;

Query Match 56.5%; Score 39; DB 16; Length 130;
Best Local Similarity 61.5%; Pred. No. 52; Mismatches 1; Mismatches 4; Indels 0; Gaps 0;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 EGGPGTTSNRLDA 13
:|||||:
DB 11 EGTGSTPRRLGA 23

RESULT 19
Q8LR30 ID Q8LR30 PRELIMINARY; PRT; 146 AA.
AC Q8LR30;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DE 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE P0671D01.21 protein.
GN P0671D01.21.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Euphorbiaceae; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartioideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1, PAC
clone: P0671D01.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF003284; BAB91744.1; -
DR Gramene; Q8LR30; -
KW SEQUENCE 146 AA; 16049 MW; CE25E9C7AA95B0B7 CRC64;

Query Match 56.5%; Score 39; DB 10; Length 146;
Best Local Similarity 70.0%; Pred. No. 58;

```

Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 3 GPGTTSNRLD 12
Db 14 GPGTSSHRSD 23

RESULT 20

Q8Q2G1 PRELIMINARY; PRT; 193 AA.
AC Q8Q2G1; (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WM1816;
RA Herring B.L., Dwyer D.E.;
RT "Australian HIV-1 env subtypes."
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY081970; AAL99288.1;
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0019031; C:Viral envelope; IEA.
DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
KW AIDS; Coat protein; Glycoprotein.
FT NON_TER 1
FT NON_TER 193
SQ SEQUENCE 193 AA; 21213 MW; 0079F7E335BAD27F CRC64;

Query Match 56.5%; Score 39; DB 15; Length 193;
Best Local Similarity 54.5%; Pred. No. 78;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EGGPGTTSNRL 11
Db 176 DGGPNSTNEI 186

RESULT 21

Q7VRN6 PRELIMINARY; PRT; 193 AA.
AC Q7VRN6;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Para-aminobenzoate synthase, glutamine amidotransferase component II (EC 4.1.3.-).
GN PABA OR BFI568.
OS Candidatus Blochmannia floridanus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; ant endosymbionts; Candidatus Blochmannia.
OX NCBI_TaxID=203907;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22784745; PubMed=12886019;
RA Gil R., Silva F.J., Zienitz E., Delmotte F., Gonzalez-Candelas F.,
RA Latorre A., Rausell C., Kamerbeek J., Gadau J., Hoelldobler B.,
RA van Ham R.C.H.J., Gross R., Moya A.;
RT "The genome sequence of Blochmannia floridanus: comparative analysis of reduced genomes."
RL Proc. Natl. Acad. Sci. U.S.A. 100:9388-9393 (2003).
DR EMBL; BX248586; CAD83250.1;
KW Glutamine amidotransferase; Lyase; Transferase; Complete proteome.
SQ SEQUENCE 193 AA; 22153 MW; 6284B168601B7356 CRC64;

Query Match 56.5%; Score 39; DB 16; Length 193;

Best Local Similarity 100.0%; Pred. No. 78;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 GPGTTSN 9
Db 54 GPGTTSN 60

RESULT 22

Q8CG89 PRELIMINARY; PRT; 278 AA.
AC Q8CG89;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Similar to LIM domains containing 1 (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CZECH II;
RA Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC042762; AAH42762.1;
DR InterPro; IPR001781; LIM.
DR Pfam; PF00412; LIM; 3.
DR ProDom; PD000094; LIM; 3.
DR SMART; SM00132; LIM; 3.
DR PROSITE; PS00478; LIM_DOMAIN_1; 2.
DR PROSITE; PS0023; LIM_DOMAIN_2; 3.
FT NON_TER 1
FT NON_TER 278
SQ SEQUENCE 278 AA; 30546 MW; E7BBBCB8F565EC28 CRC64;

Query Match 56.5%; Score 39; DB 11; Length 278;
Best Local Similarity 53.8%; Pred. No. 11e+02;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 EGGPGTTSNRLDA 13
Db 40 EGGPSAAERLEA 52

RESULT 23

Q9GLY9 PRELIMINARY; PRT; 300 AA.
AC Q9GLY9;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hypothetical protein.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA Suzuki Y., Sugano S., Hashimoto K.;
RT "Isolation of full-length cDNA clones from macaque brain CDNA libraries."
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB050521; BAB17289.1;
DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:008270; F:zinc ion binding; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR000306; Znf_FYVE.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF01363; FYVE; 1.

```

DR SMART; SM00064; FYVE; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS00190; CYTOCHROME C; 1.
DR PROSITE; PS0178; ZF_FYVE; 1.
DR Hypothetical protein.
KW Hypothetical protein.
SQ SEQUENCE 300 AA; 33123 MW; AEB590E5908795F CRC64;

Query Match          56.5%; Score 39; DB 6; Length 300;
Best Local Similarity 46.2%; Pred. No. 1.2e+02;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 EGGPGTTSNRLDA 13
DB 274 QGGPGSTNSKKQA 286
      :||||:|
      :||||:|

RESULT 24
Q98MF5 PRELIMINARY; PRT; 301 AA.
AC Q98MF5;
DT 01-OCT-2001 (TREMBlrel. 18, Created)
DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
DT 01-OCT-2001 (TREMBlrel. 25, Last annotation update)
DE Proline iminopeptidase.
GN MLL0601.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpō S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338 (2000).
DR EMBL; AF002995; BAB48158.1; -.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0016804; F:prolyl aminopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000073; A/B hydrolase.
DR InterPro; IPR002410; Peptidase S33.
DR InterPro; IPR005945; Peptidase S33.
DR Pfam; PF00561; abhydrolase_1.
DR PRINTS; PR00733; PROAMNOPTASE.
DR TIGRFAbs; TIGR01250; pro_imino_pep_2; 1.
KW Complete proteome.
SQ SEQUENCE 301 AA; 33514 MW; FD668FEFD296838F CRC64;

Query Match          56.5%; Score 39; DB 16; Length 301;
Best Local Similarity 58.3%; Pred. No. 1.2e+02;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 GGGPGTTSNRLDA 13
DB 39 GGGGAHNYVDA 50
      :||||:|
      :||||:|

RESULT 25
Q98VF5 PRELIMINARY; PRT; 302 AA.
AC Q98VF5;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
DE CG13727 protein.
CG13727.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blake J.R.G., Champ M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abrell J.F., Agbayani A., An H.-J., Andrews-Frankoch C., Baldwin D.,
RA Balles R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glaeser K.,
RA Glöckle A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Matzel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195 (2000).
DR EMBL; AE003524; AAF49356.1; -.
DR FlyBase; FBgn0036711; CG13727.
SQ SEQUENCE 302 AA; 31787 MW; 5D373F9C24329831 CRC64;

Query Match          56.5%; Score 39; DB 5; Length 302;
Best Local Similarity 54.5%; Pred. No. 1.2e+02;
Matches 6; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 3 GGGTTSNRLDA 13
DB 71 GPGSSSKLDS 81
      :||||:|
      :||||:|

Search completed: September 24, 2004, 02:18:54
Job time : 69.3019 secs

```

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2005, 17:57:45 ; Search time 84.0938 Seconds
(without alignments)

79.162 Million cell updates/sec

Title: US-10-088-639A-2_COPY_226_238

Perfect score: 69

Sequence: 1 EGGPGTTSNRLDA 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt_03:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44	63.8	419	2 Q7UNR6	Q7unr6 rhodopirell
2	44	63.8	3925	1 BSN_HUMAN	Q9upa5 homo sapien
3	43	62.3	148	2 Q62L93	Q62l93 burkholderi
4	43	62.3	148	2 Q63V99	Q63v99 burkholderi
5	43	62.3	149	1 RNH_CAUCR	Q9a341 caulobacter
6	43	62.3	749	2 Q9Y149	Q9y149 drosophila
7	42	60.9	313	2 Q89GJ0	Q89gj0 bradyrhizob
8	42	60.9	503	2 Q8BXM8	Q8bxm8 mus musculu
9	42	60.9	541	2 Q8Q2T7	Q8q2t7 mus musculu
10	42	60.9	579	2 Q8C0G3	Q8c0g3 mus musculu
11	42	60.9	588	2 Q8C078	Q8c078 mus musculu
12	42	60.9	588	2 Q8CH42	Q8ch42 mus musculu
13	42	60.9	1513	2 Q90Z70	Q90z70 brachydanio
14	41	59.4	148	2 Q63V77	Q63v77 burkholderi
15	41	59.4	166	2 Q62EP2	Q62ep2 burkholderi
16	40	58.0	105	2 Q8RPU6	Q8wpu6 trypanosoma
17	40	58.0	229	1 GL12_ARATH	Q9m8x1 arabidopsis
18	40	58.0	249	2 Q7BW10	Q7bw10 salmonella
19	40	58.0	249	2 Q84893	Q84893 salmonella
20	40	58.0	283	2 Q9XSZ8	Q9xsz8 cercopithec
21	40	58.0	313	2 Q9RL33	Q9rl33 streptomyce
22	40	58.0	344	1 DHBI_RAT	P51657 rattus norv
23	40	58.0	351	2 Q9XAAB8	Q9xab8 streptomyce
24	40	58.0	757	2 Q82WD2	Q82wd2 nitrosomona
25	40	58.0	937	2 Q7SF69	Q7sf69 neurospora
26	40	58.0	1106	2 Q7R1P8	Q7r1p8 giardia lam
27	39	56.5	67	2 Q92MH7	Q92mh7 rhizobium m
28	39	56.5	130	2 Q883M3	Q883m3 pseudomonas
29	39	56.5	146	2 Q8LR30	Q8lr30 oryza sativ
30	39	56.5	193	2 Q7VRN6	Q7vrn6 candidatus
31	39	56.5	193	2 Q8Q2G1	Q8q2g1 human immun

32	39	56.5	278	2 Q8CG89	Q8cgg89 mus musculu
33	39	56.5	300	2 Q9GLY9	Q9gly9 macaca faec
34	39	56.5	301	2 Q98MF5	Q98mf5 rhizobium l
35	39	56.5	302	2 Q9VVF5	Q9vvf5 drosophila
36	39	56.5	385	2 Q8LZW8	Q8lzw8 streptomyce
37	39	56.5	396	2 Q96K21	Q96k21 homo sapien
38	39	56.5	398	2 Q7TQJ8	Q7tqj8 mus musculu
39	39	56.5	422	2 Q9SSE1	Q9sse1 drosophila
40	39	56.5	454	2 Q834L7	Q834l7 enterococcu
41	39	56.5	473	2 Q84WQ1	Q84wq1 arabidopsis
42	39	56.5	499	2 Q85WC2	Q85wc2 homo sapien
43	39	56.5	521	2 Q6INA0	Q6ina0 homo sapien
44	39	56.5	615	2 Q9KX0	Q9kx0 streptomyce
45	39	56.5	750	2 Q8BVV4	Q8bv4 mus musculu

ALIGNMENTS

RESULT 1
Q7UNR6 PRELIMINARY; PRT; 419 AA.
AC Q7UNR6
DT 01-OCT-2003 (Tremblrel. 25, Created)
DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Amino-transferase Nifs-putative Nifs homolog (EC 4.4.1.-).
GN Name=nifs; OrderedLocusNames=RB7424;
OS Rhodopirellula baltica.
OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Pirellula.
OX NCBI_TaxID=117;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1;
RX MEDLINE=22735913; PubMed=12835416; DOI=10.1073/pnas.1431443100;
RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,
RA Schlesner H., Amann R., Reinhardt R.;
RT "Complete genome sequence of the marine planctomycete Pirellula sp.
RL strain 1.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
DR EMBL; BX294146; CAD75352.1; -.
DR HSSP; P77444; 1JF9.
DR GO; GO:0016829; P:lyase activity; IEA.
DR GO; GO:0008483; P:transaminase activity; IEA.
DR GO; GO:0016740; P:transferase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR000192; Aminotrans V.
DR Pfam; PF00266; Aminotran 5; 1.
KW Amino-transferase; Complete proteome; Lyase; Transferase.
SQ SEQUENCE 419 AA, 43485 MW, B91E67E8F2C62D79 CRC64;

Query Match 63.8%; Score 44; DB 2; Length 419;
Best Local Similarity 66.7%; Pred. No. 44;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 GGGTTSNRLDA 13
|||
DB 392 GGGTTSNRLDA 403

RESULT 2
BSN_HUMAN STANDARD; PRT; 3925 AA.
ID BSN_HUMAN
AC Q9UPA5; O43161; Q7LGH3;
DT 25-OCT-2004 (Rel. 45, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Basoon protein (Zinc-finger protein 231).
GN Name=BSN; Synonyms=KIAA0434, ZNF231;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RX MEDLINE=99026125; PubMed=9806829; DOI=10.1006/geno.1998.5516;
RA Hashida H., Goto J., Zhao J., Zhao N., Takahashi N., Hirai M., Kanazawa I.,
RA Sakaki Y.;
RT "Cloning and mapping of ZNF231, a novel brain-specific gene encoding
RT neuronal double zinc finger protein whose expression is enhanced in a
RT neurodegenerative disorder, multiple system atrophy.";
RL Genomics 54:50-58(1998).
RN [2]
RP SEQUENCE OF 76-3925 FROM N.A.
RX PubMed=10329005; DOI=10.1006/geno.1999.5788;
RA Winter C., tom Dieck S., Boeckers T., Bockmann J., Kaempf U.,
RA Sanmarti-Vila L., Langnaese K., Altmann W., Stumm M., Soyke A.,
RA Wieacker P., Garner C.C., Gundelfinger E.D.;
RT "The presynaptic cytomatrix protein Bassoon: sequence and chromosomal
RT localization of the human BSN gene.";
RL Genomics 57:389-397(1999).
RN [3]
RP SEQUENCE OF 2356-3925 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98116655; PubMed=9455477;
RA Ishikawa K.-I., Nagase T., Nakajima D., Seki N., Ohira M.,
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. VIII.
RT 78 new cDNA clones from brain which code for large proteins in
RT vitro.";
RL DNA Res. 4:307-313(1997).
CC -!- FUNCTION: Is thought to be involved in the organization of the
CC cytomatrix at the nerve terminals active zone (CAZ) which
CC regulates neurotransmitter release. Seems to act through binding
CC to ERC2/CAS1. Essential in regulated neurotransmitter release
CC from a subset of brain glutamatergic synapses. Involved in the
CC formation of the retinal photoreceptor ribbon synapses (By
CC similarity).
CC -!- SUBUNIT: Interacts with ERC2/CAS1, RIMS1 and UNC13A. Part of a
CC complex consisting of ERC2, RIMS1 and BSN (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic; localized to the active zone of
CC presynaptic density (By similarity).
CC -!- TISSUE SPECIFICITY: Exclusively expressed in brain.
CC -!- PTM: Myristoylated. The N-terminal myristoylation is not
CC sufficient for presynaptic localization (By similarity).
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF052224; AAC83555.1; --
DR EMBL: Y18448; CAA77176.1; --
DR EMBL: Y18449; CAA77176.1; JOINED.
DR EMBL: Y18450; CAA77176.1; JOINED.
DR EMBL: Y18451; CAA77176.1; JOINED.
DR EMBL: AB007894; BAA23707.1; --
DR FIC: T00062; T00062.
DR Genbank: HGNC:1117; BSN.
DR MIM: 604020; --
DR GO: GO:0005634; C:nucleus; TAS.
DR GO: GO:0007268; P:synaptic transmission; TAS.
DR InterPro: IPR011011; FYVE PHD ZNF.
DR InterPro: IPR008899; Znf_Piccolo.
DR Pfam: PF05715; zf_piccolo; 2.
KW Coiled coil; Cytochrome; Lipoprotein; Metal-binding; Myristate;
KW Repeat; Synapse; Synaptosome; Zinc; Zinc-finger.
FT INIT MET 0 0 By similarity.
FT LIPID 1 1 N-myristoyl glycine (By similarity).
FT DOMAIN 22 31 5 x 2 AA tandem repeats of P-G.
FT DOMAIN 60 73 7 x 2 AA tandem repeats of P-G.

FT ZN_FING 169 192 C4-type (Potential).
FT ZN_FING 197 219 C4-type (Potential).
FT ZN_FING 464 487 C4-type (Potential).
FT ZN_FING 492 514 C4-type (Potential).
FT DOMAIN 570 590 3 x 7 AA tandem repeats of K-A-S-P-[LQ] -
FT [APS] - [KST].
FT REPEAT 570 576 1.
FT REPEAT 577 583 2.
FT REPEAT 584 590 3.
FT DOMAIN 1036 1091 Coiled coil (Potential).
FT DOMAIN 1180 1207 Coiled coil (Potential).
FT DOMAIN 1275 1293 Coiled coil (Potential).
FT DOMAIN 1350 1475 Coiled coil (Potential).
FT DOMAIN 2938 2980 Coiled coil (Potential).
FT DOMAIN 2999 2605 Poly-Arg.
FT DOMAIN 2626 2631 Poly-Arg.
FT DOMAIN 3774 3782 Poly-Gln.
FT CONFLICT 3924 3924 F -> L (in Ref. 1).
SQ SEQUENCE 3925 AA; 416365 MW; 38A910B2BCACC3A CRC64;

Query Match 63.8%; Score 44; DB 1; Length 3925;
Best Local Similarity 80.0%; Pred. No. 5.6e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 GPGTTSNRLLD 12
DB 71 GPGSTSRRLD 80

RESULT 3
Q62L93 PRELIMINARY; PRT; 148 AA.
AC Q62L93;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Ribonuclease HI (EC 3.1.26.4).
GN Name=rRNA; ORFNames=EMAO763;
OS Burkholderia mallei ATCC 23344.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia.
OX NCBI_TaxID=243160;
RN [1]_TaxID=243160;
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 23344;
RA Nierman W.C., DeShazer D., Kim H.S., Tettelin H., Nelson K.E.,
RA Feldblyum T., Ulrich R.L., Ronning C.M., Brinkac L.M., Daugherty S.C.,
RA Davidson T.D., Deboy R.T., Dimitrov G., Dodson R.J., Durkin A.S.,
RA Gwinn M.L., Haft D.H., Khouri H., Kolonay J.F., Madupu R.,
RA Mohammad Y., Nelson W.C., Radune D., Romero C.M., Sarria S.,
RA Selengut J., Shambhlin C., Sullivan S.A., White O., Yu Y., Zafar N.,
RA Zhou L., Fraser C.M.;
RT "Structural flexibility in the Burkholderia mallei genome.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:14247-14251(2004).
DR EMBL: CP000010; AAU49148.1; --
KW Hydrolase.
SQ SEQUENCE 148 AA; 16254 MW; C90D92F92CB1B60C CRC64;

Query Match 62.3%; Score 43; DB 2; Length 148;
Best Local Similarity 63.6%; Pred. No. 20;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 GPGTTSNRLLD 12
DB 39 GPGTTSNRRLD 49

RESULT 4
Q63V99 PRELIMINARY; PRT; 148 AA.
AC Q63V99;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)

25-OCT-2004 (TReMBLrel. 28, Last annotation update)
CC Ribonuclease HI (EC 3.1.26.4).
CC GN Name=rnhA; Synonyms=dasF, herA, rnh, sdrA; ORFNames=BP6L1342;
OS Burkholderia pseudomallei K96243
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
CC Burkholderiaceae; Burkholderia.
CC NCBI_TaxID=272560;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K96243;
RX PubMed=15377794;
RA Holden M.T.G., Titball R.W., Peacock S.J., Cerdano-Tarraga A.M.,
RA Atkins T., Crossman L.C., Pitt T., Churcher C., Mungall K.,
RA Bentley S.D., Sebaihia M., Thomson N.R., Bason N., Beacham I.R.,
RA Brooks K., Brown K.A., Brown N.F., Challis G.L., Cherevach I.,
RA Chillingworth T., Cronin A., Crosset B., Davis P., Deshazer D.,
RA Fellwell T., Fraser A., Hance Z., Hauser H., Holroyd S., Jagels K.,
RA Keith K.E., Maddison M., Moule S., Price C., Quail M.A.,
RA Rabinowitch E., Rutherford K., Sanders M., Simmonds M.,
RA Songisvilai S., Stevens K., Tumapa S., Vesaratchavest M.,
RA Whitehead S., Yeats C., Barrell B.G., Oyston P.C.F., Parkhill J.;
RT "Genomic plasticity of the causative agent of melioidosis,
RT Burkholderia pseudomallei";
RL Proc. Natl. Acad. Sci. U.S.A. 101:14240-14245 (2004).
KW Hydrolase.
DR EMBL; BX571965; CAH35340.1; --
SQ SEQUENCE 148 AA; 16254 MW; C90D92F92CB1B60C CRC64;

Query Match 62.3%; Score 43; DB 2; Length 148;
Best Local Similarity 63.6%; Pred. No. 20;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 GCGGTTSNRDL 12
DB | |||||:::
39 GCGGTNNRME 49

RESULT 5
RHH CAUCR STANDARD; PRT; 149 AA.
ID _RNH CAUCR
AC Q9A341;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Ribonuclease HI (EC 3.1.26.4) (RNase HI).
GN Name=rnhA; OrderedLocusNames=CC3365;
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacteriales;
CC Caulobacteraceae; Caulobacter.
CC NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647; DOI=10.1073/pnas.061029298;
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J.A., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Fotocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA DeBoy R.J., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H.M., Shetty J.,
RA Berry K.J., Utterback T.R., Tran K., Wolf A.M., Vamathevan J.J.,
RA Ermolaeva M.D., White O., Salzberg S.L., Venter J.C., Shapiro L.,
RA Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141 (2001).
CC -!- FUNCTION: This enzyme is an endonuclease that degrades the RNA of
CC RNA-DNA hybrids specifically (By similarity).
CC -!- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphonate.
CC -!- COFACTOR: Binds 1 magnesium ion per subunit (By similarity).
CC -!- SUBUNIT: Monomer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -!- SIMILARITY: Belongs to the RNase H family.
CC -!- SIMILARITY: Contains 1 RNase H domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL; AE005998; AAK25327.1; --
DR PIR; C87666; C87666.
DR HSSP; P00647; IG15.
DR TIGR; CC3365; --
DR HAMAP; MF 00042; --; 1.
DR InterPro; IPR002156; RNaseH.
DR Pfam; PF00075; RNaseH; 1.
DR PROSITE; PS50879; RNaseH; 1.
KW Complete proteome; Endonuclease; Hydrolase; Magnesium; Nuclease.
FT DOMAIN 1 142 RNase H.
FT METAL 10 10 Magnesium (By similarity).
FT METAL 48 48 Magnesium (By similarity).
FT METAL 70 70 Magnesium (By similarity).
FT METAL 134 134 Magnesium (By similarity).
SQ SEQUENCE 149 AA; 16674 MW; CB2271A33782435C CRC64;

Query Match 62.3%; Score 43; DB 1; Length 149;
Best Local Similarity 63.6%; Pred. No. 21;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 GCGGTTSNRDL 12
DB | |||||:::
38 GCGGTNNRME 48

RESULT 6
QY149 PRELIMINARY; PRT; 749 AA.
ID QY149
AC QY149;
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
DE CG4184-PA (BcdNA.GH03922).
GN Name=Arx105; Synonyms=BcdNA.GH03922; ORFNames=CG4184;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
CC NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers J.H., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Aghayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Gaig N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpun G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merklov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodgett, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195 (2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=2426065; PubMed=12537558;
 RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
 RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
 RA George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,
 RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
 RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
 RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
 RT "Finishing a whole-genome shotgun: Release 3 of the *Drosophila*
 RT *melanogaster* euchromatic genome sequence.";
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=2426070; PubMed=12537573;
 RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
 RA Ashburner M., Celniker S.E.;
 RT "The transposable elements of the *Drosophila melanogaster* euchromatin:
 RT a genomics perspective.";
 RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=2426069; PubMed=12537572;
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.B.,
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktarglu L., Berman B.P.,
 RA Bettencourt B.K., Celniker S.E., de Grey A.D., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.;
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a
 RT systematic review";
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
 RN [5]
 RP SEQUENCE FROM N.A.
 RG FlyBase;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RG FlyBase;
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE FROM N.A.
 RA Rubin G.M., Wan K.H., Harvey D., Lewis S.E., Brokstein P., Tsang G.,
 RA Agbayani A., Arcaina T.T., Baxter E., Blazej R.G., Butenhoff C.,
 RA Champe M., Chavez C., Chew M., Doyle C.M., Farfan D.E., Frise E.,
 RA Galle R., George R.A., Harris N.L., Hoskins R.A., Evans-Holm M.,
 RA Houston K.A., Hummasti S.R., Kim E., Li P., Moshrefi M., Pacleb J.M.,
 RA Park S., Sequeira A., Sethi H., Shier E., Svirskas R.R., Weinburg T.,
 RA Celniker S.E.;
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF5003589; AAF51490.1;
 DR EMBL; AF145620; AAD38595.1;
 DR FlyBase; Fggn0027592; Arc105.
 SQ SEQUENCE 749 AA; 80526 MW; 42375AEB7C13A1ADF CRC64;

Query Match 62.3%; Score 43; DB 2; Length 749;
 Best Local Similarity 66.7%; Pred. No. 1.3e+02;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 2 GGGPGTTSNRIDA 13
 DB 127 GGGPGTASNLQSS 138
 ID Q89GJ0 PRELIMINARY; PRT; 313 AA.
 AC Q89GJ0;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE ABC transporter permease protein.
 GN OrderedLocustNames=bl16355;
 OS Bradyrhizobium japonicum.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Bradyrhizobiaceae; Bradyrhizobium.
 OX NCBI_TaxID=375;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=USD110;
 RX MEDLINE=22484998; PubMed=12597275;
 RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiomi T.,
 RA Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
 RA Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,
 RA Tabata S.;
 RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
 RT Bradyrhizobium japonicum USD110.";
 RL DNA Res. 9:189-197(2002).
 CC -!- FUNCTION: Part of a binding-protein-dependent transport system.
 CC Probably responsible for the translocation of the substrate across
 CC the membrane (By similarity).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
 CC -!- SIMILARITY: Belongs to the binding-protein-dependent transport
 CC system permease family.
 DR EMBL; AP005958; BAC51620.1;
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005215; F:transporter activity; IEA.
 DR GO; GO:0008810; P:transport; IEA.
 DR InterPro; IPR000515; BPD transp.
 DR Pfam; PF00528; BPD_transp_1; 1.
 DR PROSITE; P850928; ABC_TM1; 1.
 KW Complete proteome; Transmembrane; Transport.
 SQ SEQUENCE 313 AA; 34112 MW; AE97D4548B8394A0 CRC64;
 Query Match 60.9%; Score 42; DB 2; Length 313;
 Best Local Similarity 58.3%; Pred. No. 71;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 1 EGGPGTTSNRILD 12
 DB 259 QGGPGNSNTLN 270
 ID Q8BXM8 PRELIMINARY; PRT; 503 AA.
 AC Q8BXM8;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Mus musculus adult retina cDNA, RIKEN full-length enriched library,
 DE clone: A93003018 product: CA+/CALMODULIN-DEPENDENT PROTEIN KINASE
 DE KINASE BETA (CAM-KINASE KINASE BETA) homolog.
 GN Name=Cankk2;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;

RN SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Retina;
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
 RA Carninci P., Hayaehizaki Y.;
 RT "High-efficiency full-length cDNA cloning.";
 RL Meth. Enzymol. 303:19-44 (1999).
 RN [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Retina;
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
 RA RIKEN FANTOM Consortium;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690 (2001).
 RN [3]
 RN SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Retina;
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
 Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
 Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
 Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RA "RIKEN integrated sequence analysis (RISA) system-384-format
 sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1757-1771 (2000).
 RN [5]
 RN SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Retina;
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
 Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
 Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
 Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
 Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RA "RIKEN integrated sequence analysis (RISA) system-384-format
 sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1757-1771 (2000).
 RN [6]
 RN SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Retina;
 RX Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
 Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
 Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
 Hori F., Inotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
 Kato H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
 Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
 Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
 Saito R., Saitoh H., Sakai K., Sakai K., Sakazume N., Sano H.,
 Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
 Tagawa A., Takahashi F., Takaku-Akahira S., Tanaka Y., Tanaka T.,
 Tomaru A., Toyota T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
 RA Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
 DR EMBL; AK044660; BAC32023.1; -.
 DR HSSP; O14965; 1OL6.
 DR MGD; MGI:2444812; CamkK2.
 DR GO; GO:0005622; C:intracellular; ISS.
 DR GO; GO:0005509; F:calcium ion binding; ISS.
 DR GO; GO:0004685; F:calcium and calmodulin-dependent protein k. . .; ISS.
 DR GO; GO:0005516; F:calmodulin binding; ISS.
 DR GO; GO:0004713; F:protein-tyrosine kinase activity; ISS.
 DR GO; GO:0046777; P:autophosphorylation; ISS.
 DR GO; GO:0019722; P:calcium-mediated signaling; ISS.
 DR GO; GO:0000165; P:MAPKK cascade; ISS.
 DR GO; GO:0045941; P:positive regulation of transcription; ISS.

DR GO; GO:0045859; P:regulation of protein kinase activity; ISS.
 DR InterPro; IPR011009; Kinase like.
 DR InterPro; IPR000719; Prot kinase.
 DR InterPro; IPR002290; Ser Thr pkinase.
 DR InterPro; IPR008271; Ser_thr_pkin_AS.
 DR Pfam; PF00069; Pkinase; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00220; S_TK; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 503 AA; 55501 MW; 8C596D6B3649F42 CRC64;

 Query Match 60.9%; Score 42; DB 2; Length 503;
 Best Local Similarity 58.3%; Pred. No. 1.2e+02;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

 QY 1 EGGPGTTSNRLLD 12
 :||| :|||
 DB 105 QGGPASSNSSLD 116

 RESULT 9
 Q8QZT7 PRELIMINARY; PRT; 541 AA.
 AC Q8QZT7;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Calcium/calmodulin-dependent protein kinase 2, beta.
 GN Name=CamkK2;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Mix FVB/N; TISSUE=Mammary tumor;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 Brownstein M.J., Udgin T.B., Toshiyuki S., Carninci P., Prange C.,
 Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
 Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 Krzywinski M.I., Skalska U., Smailus D.E., Schnerk A., Schein J.E.,
 Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Mix FVB/N; TISSUE=Mammary tumor;
 RA Strausberg R.;
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
 DR EMBL; BC023103; RAH23103.1; -.
 DR HSSP; O14965; 1OL6.
 DR MGD; MGI:2444812; CamkK2.
 DR GO; GO:0005622; C:intracellular; ISS.
 DR GO; GO:0005509; F:calcium ion binding; ISS.
 DR GO; GO:0004685; F:calcium- and calmodulin-dependent protein k. . .; ISS.
 DR GO; GO:0005516; F:calmodulin binding; ISS.

```

DR GO: 0004713; F:protein-tyrosine kinase activity; ISS.
DR GO: 0004677; P:autophosphorylation; ISS.
DR GO: 0019722; P:calcium-mediated signaling; ISS.
DR GO: 0000165; P:MAPKK cascade; ISS.
DR GO: 0045941; P:positive regulation of transcription; ISS.
DR GO: 0045859; P:regulation of protein kinase activity; ISS.
DR InterPro: IPR011009; Kinase like.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR InterPro: IPR008271; Ser_thr_pkin_AS.
DR Pfam: PF00069; Pkinase; 1.
DR ProDom: PD000001; Prot_kinase; 1.
DR SMART: SM00220; S_TKC; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 541 AA; 59616 MW; FEB2385A4D793FF2 CRC64;

Query Match 60.9%; Score 42; DB 2; Length 541;
Best Local Similarity 58.3%; Pred. No. 1.3e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EGGPGTTSNRDL 12
DB 105 QGGPASSNSLD 116

RESULT 10
Q8C0G3 PRELIMINARY; PRT; 579 AA.
AC Q8C0G3;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Mus musculus 13 days embryo male testis cDNA, RIKEN full-length
DE enriched library, clone:6030423C20 product:CA+/CALMODULIN-DEPENDENT
DE PROTEIN KINASE KINASE BETA (CAM-KINASE KINASE BETA) homolog.
GN Name=Cankk2;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1] _TaxID=10090;
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
[2]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
[3]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RA The FANTOM Consortium;
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
[4]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata K., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes.";

```

```

RL GO: 0004713; F:protein-tyrosine kinase activity; ISS.
RN GO: 0004677; P:autophosphorylation; ISS.
RP GO: 0019722; P:calcium-mediated signaling; ISS.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
[6]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saico R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DDAJ databases.
CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL; AK031399; BAC27387.1; -.
DR HSSP; O14965; IOL6.
DR MGD; MGI:2444812; Cankk2.
DR GO: 0005622; C:intracellular; ISS.
DR GO: 0005509; F:calcium ion binding; ISS.
DR GO: 0004685; F:calcium- and calmodulin-dependent protein k. . .; ISS.
DR GO: 0005516; F:calmodulin binding; ISS.
DR GO: 0004713; F:protein-tyrosine kinase activity; ISS.
DR GO: 0046777; P:autophosphorylation; ISS.
DR GO: 0000165; P:calcium-mediated signaling; ISS.
DR GO: 00045941; P:positive regulation of transcription; ISS.
DR GO: 0045859; P:regulation of protein kinase activity; ISS.
DR InterPro: IPR011009; Kinase like.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR InterPro: IPR008271; Ser_thr_pkin_AS.
DR Pfam: PF00069; Pkinase; 1.
DR ProDom: PD000001; Prot_kinase; 1.
DR SMART: SM00220; S_TKC; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 579 AA; 63477 MW; A79C6F3E1CE6DDCA CRC64;

Query Match 60.9%; Score 42; DB 2; Length 579;
Best Local Similarity 58.3%; Pred. No. 1.4e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EGGPGTTSNRDL 12
DB 105 QGGPASSNSLD 116

RESULT 11
Q8C078 PRELIMINARY; PRT; 588 AA.
ID Q8C078;
AC Q8C078;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)

```

DE Mus musculus adult male medulla oblongata cDNA, RIKEN full-length
 DE enriched library, clone:6330570N16 product:CA+/CALMODULIN-DEPENDENT
 DE PROTEIN KINASE BETA (CAM-KINASE KINASE BETA) homolog.
 GN Name=Camk2;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Medulla oblongata;
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning.";
 RL Meth. Enzymol. 303:19-44(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Medulla oblongata;
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
 RA RIKEN FANTOM Consortium;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Medulla oblongata;
 RX MEDLINE=20499374; PubMed=11076861; DOI=10.1101/gr.145100;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";
 RL Genome Res. 10:1617-1630(2000).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Medulla oblongata;
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
 RA Konno H., Akiyama J., Nishi K., Kitsuunai T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multiplex capillary sequencer.";
 RL Genome Res. 10:1757-1771(2000).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Medulla oblongata;
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
 RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
 RA Hori F., Inotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
 RA Saio K., Saichou H., Sakai C., Sakai K., Sakazume N., Sano H.,
 RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
 RA Tagawa A., Takahashi P., Takahashi-Akai H., Tanaka Y., Tanaka T.,
 RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.,
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
 DR EMBL; AK032070; BAC27681.1; -.
 DR HSSP; O14965; 1OL6.
 DR MGD; MGI:2444812; Camk2.
 DR GO; GO:0005622; C:intracellular; ISS.

DR GO; GO:0005509; F:calcium ion binding; ISS.
 DR GO; GO:0004685; F:calcium- and calmodulin-dependent protein k...; ISS.
 DR GO; GO:0005516; F:calmodulin binding; ISS.
 DR GO; GO:0004713; F:protein-tyrosine kinase activity; ISS.
 DR GO; GO:0046777; F:autophosphorylation; ISS.
 DR GO; GO:0019722; F:calcium-mediated signaling; ISS.
 DR GO; GO:0000165; F:MAPKK cascade; ISS.
 DR GO; GO:0045941; P:positive regulation of transcription; ISS.
 DR GO; GO:0045859; P:regulation of protein kinase activity; ISS.
 DR InterPro; IPR011009; Kinase like.
 DR InterPro; IPR000719; Prot kinase.
 DR InterPro; IPR002290; Ser_thr_kinase.
 DR InterPro; IPR008271; Ser_thr_pkin_AS.
 DR Pfam; PF00069; Pkinase; 1.
 DR ProDom; PD000001; Prot kinase; 1.
 DR SMART; SM00220; S_TKG; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 588 AA; 64598 MW; 95B13E1C28FC614F CRC64;
 Query Match 60.9%; Score 42; DB 2; Length 588;
 Best Local Similarity 58.3%; Pred. No. 1.5e+02;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 1 EGGPGTTSNRLD 12
 DB 105 QGGPSSNSLD 116
 |||||:|||||
 RESULT 12
 O8CH42 PRELIMINARY; PRT; 588 AA.
 ID O8CH42
 AC O8CH42
 DT 01-WAR-2003 (TRENBLrel. 23, Created)
 DT 01-WAR-2003 (TRENBLrel. 23, Last sequence update)
 DT 01-WAR-2004 (TRENBLrel. 26, Last annotation update)
 DE Ca/calmodulin-dependent protein kinase beta.
 GN Name=Camk2;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Carra S., Harvey M., Tacedda F., Barden N.;
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
 DR EMBL; AF453383; AAN75696.1; -.
 DR HSSP; O14965; 1OL6.
 DR MGD; MGI:2444812; Camk2.
 DR GO; GO:0005622; C:intracellular; ISS.
 DR GO; GO:0005509; F:calcium ion binding; ISS.
 DR GO; GO:0004685; F:calcium- and calmodulin-dependent protein k...; ISS.
 DR GO; GO:0005516; F:calmodulin binding; ISS.
 DR GO; GO:0004713; F:protein-tyrosine kinase activity; ISS.
 DR GO; GO:0046777; F:autophosphorylation; ISS.
 DR GO; GO:0019722; F:calcium-mediated signaling; ISS.
 DR GO; GO:0000165; F:MAPKK cascade; ISS.
 DR GO; GO:0045941; P:positive regulation of transcription; ISS.
 DR GO; GO:0045859; P:regulation of protein kinase activity; ISS.
 DR InterPro; IPR011009; Kinase like.
 DR InterPro; IPR000719; Prot kinase.
 DR InterPro; IPR002290; Ser_thr_kinase.
 DR InterPro; IPR008271; Ser_thr_pkin_AS.
 DR Pfam; PF00069; Pkinase; 1.
 DR ProDom; PD000001; Prot kinase; 1.
 DR SMART; SM00220; S_TKG; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.

```

SQ SEQUENCE 588 AA; 64753 MW; 8026EE485C3BA06A CRC64;
Query Match 60.9%; Score 42; DB 2; Length 588;
Best Local Similarity 58.3%; Pred. No. 1.5e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EGGPGTTSNRDL 12
Db 105 QGGPSSNSLD 116
:||||:|||||

RESULT 13
Q90270 PRELIMINARY; PRT; 1513 AA.
AC Q90270;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Roundabout2.
GN Name=robo2;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21270059; PubMed=11376489; DOI=10.1002/dvdy.1136;
RA Lee J.-S., Ray R., Chien C.B.;
RT "Cloning and expression of three zebrafish roundabout homologs suggest
RT roles in axon guidance and cell migration.";
RL Dev. Dyn. 221:216-230(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Lee J.-S., Ray R., Chien C.-B.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF337035; AAK58427.1; -.
DR HSSP; QHCK4; IUEM.
DR ZFIN; ZDB-GENE-001019-1; robo2.
DR GO; GO:0009986; C:cell surface; ISS.
DR GO; GO:0042802; P:protein self binding; ISS.
DR GO; GO:0007420; P:brain development; ISS.
DR GO; GO:0007156; P:homophilic cell adhesion; ISS.
DR GO; GO:0050772; P:positive regulation of axonogenesis; ISS.
DR GO; GO:0001657; P:ureteric bud development; ISS.
DR Pfam; PF00041; fn3; 3.
DR SMART; SM00060; FN3; 3.
DR SMART; SM00408; IGC2; 3.
DR PROSITE; PS50853; FN3; 3.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS50835; IG_LIKE; 5.
SQ SEQUENCE 1513 AA; 165181 MW; D1743BACC089F0 CRC64;

Query Match 60.9%; Score 42; DB 2; Length 1513;
Best Local Similarity 72.7%; Pred. No. 4.2e+02;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 GPGTTSNRDLA 13
Db 1350 GPGTVNRRDA 1360
|||||

RESULT 14
Q63Y77 PRELIMINARY; PRT; 148 AA.
AC Q63Y77;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Avidin family protein.
GN ORFNames=BPSL0311;
OS Burkholderia pseudomallei K96243.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;

```

```

OC Burkholderiaceae; Burkholderia.
OX NCBI_TaxID=272560;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K96243;
RX PubMed=15377794;
RA Holden M.T.G., Titball R.W., Peacock S.J., Cerdeno-Tarraga A.M.,
RA Atkins T., Crossman D.C., Pitt T., Churcher C., Mungall K.,
RA Bentley S.D., Sebaihia M., Thomson N.R., Bason N., Beacham I.R.,
RA Brooks K., Brown K.A., Brown N.F., Challis G.L., Cherevach I.,
RA Chillingworth T., Cronin A., Crosset B., Davis P., Deshazer D.,
RA Feltwell T., Fraser A., Hance Z., Hauser H., Holroyd S., Jagels K.,
RA Keith K.E., Maddison M., Moule S., Price C., Quail M.A.,
RA Rabinowitsch E., Rutherford K., Sanders M., Simmonds M.,
RA Songvilai S., Stevens K., Tumapa S., Vesaratchavest M.,
RA Whitehead S., Yeats C., Barrell B.G., Oyston P.C.F., Parkhill J.;
RT "Genomic plasticity of the causative agent of melioidosis,
RT Burkholderia pseudomallei.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:14240-14245(2004).
DR EMBL; BX571965; CAH34300.1; -.
SQ SEQUENCE 148 AA; 15945 MW; 763EE611914EB7D7 CRC64;

Query Match 59.4%; Score 41; DB 2; Length 148;
Best Local Similarity 50.0%; Pred. No. 45;
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 EGGPGTTSNRDL 12
Db 109 QGGPGASSDRIN 120
:||||:|||||

RESULT 15
Q62EP2 PRELIMINARY; PRT; 166 AA.
AC Q62EP2;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Streptavidin, putative.
GN ORFNames=BMA3378;
OS Burkholderia mallei ATCC 23344.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia.
OX NCBI_TaxID=243160;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 23344;
RA Nierman W.C., Deshazer D., Kim H.S., Tettelin H., Nelson K.E.,
RA Feldblyum T., Ulrich R.L., Rinning C.M., Brinkac L.M., Daugherty S.C.,
RA Davidson T.D., Deboy R.T., Dimitrov G., Dodson R.J., Durkin A.S.,
RA Gwinn M.L., Haft D.H., Khouri H., Kolonay J.F., Madupu R.,
RA Mohammed Y., Nelson W.C., Radune D., Romero C.M., Sarría S.,
RA Selengut J., Shamblin C., Sullivan S.A., White O., Yu Y., Zafar N.,
RA Zhou L., Fraser C.M.;
RT "Structural flexibility in the Burkholderia mallei genome.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:14247-14251(2004).
DR EMBL; CF000010; AAU48617.1; -.
SQ SEQUENCE 166 AA; 17838 MW; 52D48F78CD5812B9 CRC64;

Query Match 59.4%; Score 41; DB 2; Length 166;
Best Local Similarity 50.0%; Pred. No. 52;
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 EGGPGTTSNRDL 12
Db 127 QGGPGASSDRIN 138
:||||:|||||

Search completed: April 28, 2005, 18:24:35
Job time : 92.0938 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 01:32:26 ; Search time 70.7736 Seconds
(without alignments)
43.915 Million cell updates/sec

Title: US-10-088-639A-2_COPY_23_33
Perfect score: 57
Sequence: 1 QGDSLKTYAS 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database : A_Geneseq_29Jan04: *
1: Genesecp1980s: *
2: Genesecp1990s: *
3: Genesecp2000s: *
4: Genesecp2001s: *
5: Genesecp2002s: *
6: Genesecp2003as: *
7: Genesecp2003bs: *
8: Genesecp2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	57	100.0	249	AA868087	Aab68087 An anti-a
2	54	94.7	104	AA893164	Aar93164 Anti-rhes
3	54	94.7	238	AA893165	Aar93165 Anti-rhes
4	53	93.0	110	AAW15538	AAW15538 Anti-TGF
5	53	93.0	253	ABP45519	ABP45519 Human Bly
6	51	89.5	242	ABP46073	ABP46073 Human Bly
7	51	89.5	243	ABP46091	ABP46091 Human Bly
8	51	89.5	247	ABP46092	ABP46092 Human Bly
9	51	89.5	248	ABP46077	ABP46077 Human Bly
10	51	89.5	249	ABP45991	ABP45991 Human Bly
11	51	89.5	250	ABP45979	ABP45979 Human Bly
12	50	87.7	11	AA95195	AA95195 Anti-plat
13	50	87.7	11	AA95216	AA95216 Anti-plat
14	50	87.7	11	AA80192	AA80192 Human gp9
15	50	87.7	101	AA822572	AA822572 Light cha
16	50	87.7	103	AA880091	AA880091 Human der
17	50	87.7	103	AA95489	AA95489 Human-der
18	50	87.7	104	AA93606	AA93606 Human ant
19	50	87.7	104	ABBO5054	ABBO5054 Fibronect
20	50	87.7	104	ABO27413	ABO27413 Anti-Rh(D
21	50	87.7	106	AAU02531	AAU02531 Anti-adip
22	50	87.7	107	AA95215	AA95215 Anti-plat
23	50	87.7	107	AA95194	AA95194 Anti-plat
24	50	87.7	107	AA82969	AA82969 Amino aci
25	50	87.7	108	AA95185	AA95185 Anti-plat

26	50	87.7	108	3	AA95179	AA95179 Anti-plat
27	50	87.7	108	3	AA95179	AA95179 Anti-plat
28	50	87.7	108	4	AA95179	AA95179 Anti-plat
29	50	87.7	108	4	AA95179	AA95179 Anti-plat
30	50	87.7	108	4	AA95179	AA95179 Anti-plat
31	50	87.7	108	4	AA95179	AA95179 Anti-plat
32	50	87.7	108	5	AA95179	AA95179 Anti-plat
33	50	87.7	108	5	AA95179	AA95179 Anti-plat
34	50	87.7	108	5	AA95179	AA95179 Anti-plat
35	50	87.7	108	5	AA95179	AA95179 Anti-plat
36	50	87.7	108	5	AA95179	AA95179 Anti-plat
37	50	87.7	108	5	AA95179	AA95179 Anti-plat
38	50	87.7	108	5	AA95179	AA95179 Anti-plat
39	50	87.7	108	5	AA95179	AA95179 Anti-plat
40	50	87.7	108	5	AA95179	AA95179 Anti-plat
41	50	87.7	108	5	AA95179	AA95179 Anti-plat
42	50	87.7	108	5	AA95179	AA95179 Anti-plat
43	50	87.7	108	5	AA95179	AA95179 Anti-plat
44	50	87.7	108	5	AA95179	AA95179 Anti-plat
45	50	87.7	108	5	AA95179	AA95179 Anti-plat
46	50	87.7	108	5	AA95179	AA95179 Anti-plat
47	50	87.7	108	5	AA95179	AA95179 Anti-plat
48	50	87.7	108	5	AA95179	AA95179 Anti-plat
49	50	87.7	108	5	AA95179	AA95179 Anti-plat
50	50	87.7	108	5	AA95179	AA95179 Anti-plat
51	50	87.7	108	5	AA95179	AA95179 Anti-plat
52	50	87.7	108	5	AA95179	AA95179 Anti-plat
53	50	87.7	108	5	AA95179	AA95179 Anti-plat
54	50	87.7	108	5	AA95179	AA95179 Anti-plat
55	50	87.7	108	5	AA95179	AA95179 Anti-plat
56	50	87.7	108	5	AA95179	AA95179 Anti-plat
57	50	87.7	108	5	AA95179	AA95179 Anti-plat
58	50	87.7	108	5	AA95179	AA95179 Anti-plat
59	50	87.7	108	5	AA95179	AA95179 Anti-plat
60	50	87.7	108	5	AA95179	AA95179 Anti-plat
61	50	87.7	108	5	AA95179	AA95179 Anti-plat
62	50	87.7	108	5	AA95179	AA95179 Anti-plat
63	50	87.7	108	5	AA95179	AA95179 Anti-plat
64	50	87.7	108	5	AA95179	AA95179 Anti-plat
65	50	87.7	108	5	AA95179	AA95179 Anti-plat
66	50	87.7	108	5	AA95179	AA95179 Anti-plat
67	50	87.7	108	5	AA95179	AA95179 Anti-plat
68	50	87.7	108	5	AA95179	AA95179 Anti-plat
69	50	87.7	108	5	AA95179	AA95179 Anti-plat
70	50	87.7	108	5	AA95179	AA95179 Anti-plat
71	50	87.7	108	5	AA95179	AA95179 Anti-plat
72	50	87.7	108	5	AA95179	AA95179 Anti-plat
73	50	87.7	108	5	AA95179	AA95179 Anti-plat
74	50	87.7	108	5	AA95179	AA95179 Anti-plat
75	50	87.7	108	5	AA95179	AA95179 Anti-plat
76	50	87.7	108	5	AA95179	AA95179 Anti-plat
77	50	87.7	108	5	AA95179	AA95179 Anti-plat
78	50	87.7	108	5	AA95179	AA95179 Anti-plat
79	50	87.7	108	5	AA95179	AA95179 Anti-plat
80	50	87.7	108	5	AA95179	AA95179 Anti-plat
81	50	87.7	108	5	AA95179	AA95179 Anti-plat
82	50	87.7	108	5	AA95179	AA95179 Anti-plat
83	50	87.7	108	5	AA95179	AA95179 Anti-plat
84	50	87.7	108	5	AA95179	AA95179 Anti-plat
85	50	87.7	108	5	AA95179	AA95179 Anti-plat
86	50	87.7	108	5	AA95179	AA95179 Anti-plat
87	50	87.7	108	5	AA95179	AA95179 Anti-plat
88	50	87.7	108	5	AA95179	AA95179 Anti-plat
89	50	87.7	108	5	AA95179	AA95179 Anti-plat
90	50	87.7	108	5	AA95179	AA95179 Anti-plat
91	50	87.7	108	5	AA95179	AA95179 Anti-plat
92	50	87.7	108	5	AA95179	AA95179 Anti-plat
93	50	87.7	108	5	AA95179	AA95179 Anti-plat
94	50	87.7	108	5	AA95179	AA95179 Anti-plat
95	50	87.7	108	5	AA95179	AA95179 Anti-plat
96	50	87.7	108	5	AA95179	AA95179 Anti-plat
97	50	87.7	108	5	AA95179	AA95179 Anti-plat
98	50	87.7	108	5	AA95179	AA95179 Anti-plat

99	50	87.7	242	5	ABP45751	Abp45751 Human Bly
100	50	87.7	242	5	ABP46035	Abp46035 Human Bly
101	50	87.7	242	5	ABP45433	Abp45433 Human Bly
102	50	87.7	242	5	ABP46010	Abp46010 Human Bly
103	50	87.7	242	5	ABP45413	Abp45413 Human Bly
104	50	87.7	242	5	ABP45422	Abp45422 Human Bly
105	50	87.7	242	5	ABP46095	Abp46095 Human Bly
106	50	87.7	243	5	ABP46045	Abp46045 Human Bly
107	50	87.7	243	5	ABP45929	Abp45929 Human Bly
108	50	87.7	243	5	ABP45936	Abp45936 Human Bly
109	50	87.7	243	5	ABP46003	Abp46003 Human Bly
110	50	87.7	243	5	ABP45931	Abp45931 Human Bly
111	50	87.7	243	5	ABP45933	Abp45933 Human Bly
112	50	87.7	243	5	ABP46096	Abp46096 Human Bly
113	50	87.7	243	5	ABP45918	Abp45918 Human Bly
114	50	87.7	244	5	ABP45913	Abp45913 Human Bly
115	50	87.7	244	5	ABP45683	Abp45683 Human Bly
116	50	87.7	244	5	ABP45681	Abp45681 Human Bly
117	50	87.7	244	5	ABP46047	Abp46047 Human Bly
118	50	87.7	244	6	AAO31139	Aao31139 Human CMO
119	50	87.7	244	6	AAO31136	Aao31136 Human CMO
120	50	87.7	245	5	ABP45908	Abp45908 Human Bly
121	50	87.7	245	5	ABP45604	Abp45604 Human Bly
122	50	87.7	245	5	ABP45815	Abp45815 Human Bly
123	50	87.7	245	6	AAO31137	Aao31137 Human CMO
124	50	87.7	246	5	ABG78329	Abg78329 Human Fv
125	50	87.7	246	5	ABP45807	Abp45807 Human Bly
126	50	87.7	246	5	ABP45691	Abp45691 Human Bly
127	50	87.7	246	5	ABP45964	Abp45964 Human Bly
128	50	87.7	246	5	ABP45289	Abp45289 Human Bly
129	50	87.7	246	5	ABP45415	Abp45415 Human Bly
130	50	87.7	246	5	ABP45181	Abp45181 Human Bly
131	50	87.7	246	5	ABP44917	Abp44917 Human Bly
132	50	87.7	246	5	ABP45297	Abp45297 Human Bly
133	50	87.7	246	5	ABP45644	Abp45644 Human Bly
134	50	87.7	246	5	ABG92026	Abg92026 Antibody
135	50	87.7	247	5	ABP45317	Abp45317 Human Bly
136	50	87.7	247	5	ABP44913	Abp44913 Human Bly
137	50	87.7	247	5	ABP44916	Abp44916 Human Bly
138	50	87.7	247	5	ABP45692	Abp45692 Human Bly
139	50	87.7	247	5	ABP45459	Abp45459 Human Bly
140	50	87.7	247	5	ABP44937	Abp44937 Human Bly
141	50	87.7	247	5	ABP44985	Abp44985 Human Bly
142	50	87.7	247	5	ABP45640	Abp45640 Human Bly
143	50	87.7	247	5	ABP45311	Abp45311 Human Bly
144	50	87.7	247	5	ABP45715	Abp45715 Human Bly
145	50	87.7	247	5	ABP45928	Abp45928 Human Bly
146	50	87.7	247	5	ABP45753	Abp45753 Human Bly
147	50	87.7	247	5	ABP45373	Abp45373 Human Bly
148	50	87.7	247	5	ABP45101	Abp45101 Human Bly
149	50	87.7	248	5	ABP44336	Abp44336 Human Bly
150	50	87.7	248	5	ABP45716	Abp45716 Human Bly

ALIGNMENTS

*RESULT 1
 AAB68087
 ID AAB68087 standard; protein; 249 AA.
 XX
 AC AAB68087;
 XX
 DT 09-JUL-2001 (first entry)
 XX
 DE An anti-alpha6beta4 integrin light chain linked to a heavy chain.
 XX
 KW Gastrointestinal epithelial tumour cell; alpha6beta4 integrin;
 KW tumour-associated antigen; metastatic disease; malignant disease;
 KW tumour typing; tumour screening; tumour.
 XX
 OS Synthetic.
 OS Macaca fascicularis.

Query Match 100.0%; Score 57; DB 4; Length 249;
 Best Local Similarity 100.0%; Pred. NO. 0.026;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QGDSLKTYVAS 11
 DB 23 QGDSLKTYVAS 33
 RESULT 2
 AAR93164
 ID AAR93164 standard; protein; 104 AA.

Location/Qualifiers
 23..33
 /note= "Complementarity determining region (CDR) 1 of the light chain"
 49..55
 /note= "Complementarity determining region (CDR) 2 of the light chain"
 88..98
 /note= "Complementarity determining region (CDR) 3 of the light chain"
 110..127
 /note= "linker"
 158..162
 /note= "Complementarity determining region (CDR) 1 of the heavy chain"
 177..193
 /note= "Complementarity determining region (CDR) 1 of the heavy chain"
 226..238
 /note= "Complementarity determining region (CDR) 1 of the heavy chain"
 WO200130854-A2.
 03-MAY-2001.
 26-OCT-2000; 2000WO-SE002082.
 28-OCT-1999; 99SE-00003895.
 (ACTI-) ACTIVE BIOTECH AB.
 Brodin TN, Karlstroem FJ, Ohlsson LG, Tordsson MJ, Kearney PP;
 Nilsson BHK;
 WPI; 2001-308619/32.
 N-PSDB; AAF84797.
 Novel antibody for diagnosis, treatment of human metastatic and malignant diseases, has binding structure for target structure displayed on cell surface of human gastrointestinal epithelial tumor and normal cells.
 Claim 1; Page 55-56; 75pp; English.
 The present sequence represents a Monkey antibody light chain linked to a heavy chain. The antibody binds to a target structure displayed in and on the cell surface of human gastrointestinal epithelial tumour cells and in a subpopulation of normal human gastrointestinal epithelial cells. The target structure especially comprises alpha6beta4 integrin. This is a tumour-associated antigen. The antibody, and its fragments, are useful for treating conditions based on anti-angiogenic mechanism in humans. They are useful for treating human metastatic and malignant disease, for in vitro, in vivo diagnosis and prognosis of human malignant disease, comprising tumour typing, tumour screening, tumour diagnosis and prognosis and monitoring premalignant conditions. Quantitative in vivo diagnosis is carried out by determining the localization of antibody to tumour deposits in humans

XX AAR93164;
 XX 29-OCT-1996 (first entry)
 XX Anti-rhesus D monoclonal antibody D7C2 light chain V region.
 XX Human monoclonal antibody; immunoglobulin isotype IgM; agglutination;
 KW rhesus positive; rhesus negative; haemolysis; lambda light chain;
 KW variable region; insect host cell; baculovirus; recombinant production.
 XX Homo sapiens.
 OS
 XX
 FH Key Location/Qualifiers
 FT Region 23..33
 FT /label= CDR1
 FT /notes= "complementarity determining region"
 FT Region 49..55
 FT /label= CDR2
 FT /notes= "complementarity determining region"
 FT Region 88..93
 FT /label= CDR3
 FT /note= "complementarity determining region"
 XX
 PN FR2724182-A1.
 XX
 PD 08-MAR-1996.
 XX
 PD 02-SEP-1994; 94FR-00010566.
 XX
 PD 02-SEP-1994; 94FR-00010566.
 XX
 PA (INSP) INST PASTEUR.
 PA (PROT-) PROTEINE PERFORMANCE.
 XX
 PI Edelman L, Margairitte C, Kaczorek M, Chaabihi H;
 XX
 DR WPI; 1996-162018/17.
 DR N-PSDB; AAT26869.
 XX
 PT Recombinant anti-rhesus D monoclonal antibody - expressed by baculovirus-
 PT transformed insect cells and useful for preventing haemolysis in new-born
 PT babies.
 XX
 PS Example 1; Page 30; 46pp; French.
 XX
 CC The human monoclonal antibody D7C2, of isotype IgM, recognises a 30-32 kD
 CC polypeptide on the membrane of rhesus positive red blood cells. The
 CC antibody agglutinates rhesus positive cells but not rhesus negative cells
 CC and is useful diagnostically and also for preventing haemolysis in new-
 CC born rhesus positive babies. Recombinant IGM-D7C2 can be produced by
 CC insect cells which have been transformed by a baculoviral vector
 CC comprising a D7C2 expression cassette. The present sequence is that of
 CC the variable region of the IGM-D7C2 light chain
 XX
 SQ Sequence 104 AA;
 Query Match 94.7%; Score 54; DB 2; Length 104;
 Best Local Similarity 90.9%; Pred. No. 0.036;
 Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QGDSLKTYVAS 11
 Db 23 QGDSLKTYVAS 33
 RESULT 3
 AAR93165
 ID AAR93165 standard; protein; 238 AA.
 XX
 AC AAR93165;
 XX
 DT 30-OCT-1996 (first entry)

XX Anti-rhesus D recombinant antibody D7C2 light chain.
 XX Human monoclonal antibody; immunoglobulin isotype IgM; agglutination;
 KW rhesus positive; rhesus negative; haemolysis; lambda light chain;
 KW variable region; insect host cell; baculovirus; recombinant production.
 XX Homo sapiens.
 OS
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..19
 FT /label= signal_peptide
 FT /note= "encoded by synthetic linker"
 FT Protein 20..238
 FT /label= light chain
 FT /note= "human lambda light chain constant region and the
 FT variable region from anti-rhesus D antibody D7C2"
 XX
 PN FR2724182-A1.
 XX
 PD 08-MAR-1996.
 XX
 PD 02-SEP-1994; 94FR-00010566.
 XX
 PD 02-SEP-1994; 94FR-00010566.
 XX
 PA (INSP) INST PASTEUR.
 PA (PROT-) PROTEINE PERFORMANCE.
 XX
 PI Edelman L, Margairitte C, Kaczorek M, Chaabihi H;
 XX
 DR WPI; 1996-162018/17.
 DR N-PSDB; AAT26888.
 XX
 PT Recombinant anti-rhesus D monoclonal antibody - expressed by baculovirus-
 PT transformed insect cells and useful for preventing haemolysis in new-born
 PT babies.
 XX
 PS Example 2; Page 33-34; 46pp; French.
 XX
 CC The human monoclonal antibody D7C2, of isotype IgM, recognises a 30-32 kD
 CC polypeptide on the membrane of rhesus positive red blood cells. The
 CC antibody agglutinates rhesus positive cells but not rhesus negative cells
 CC and is useful diagnostically and also for preventing haemolysis in new-
 CC born rhesus positive babies. Recombinant IGM-D7C2 can be produced by
 CC insect cells which have been transformed by a baculoviral vector
 CC comprising a D7C2 expression cassette. The present sequence is that of a
 CC recombinant IGM-D7C2 lambda light chain fused downstream of a signal
 CC peptide
 XX
 SQ Sequence 238 AA;
 Query Match 94.7%; Score 54; DB 2; Length 238;
 Best Local Similarity 90.9%; Pred. No. 0.089;
 Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QGDSLKTYVAS 11
 Db 42 QGDSLRTYYAS 52
 RESULT 4
 AAW15538
 ID AAW15538 standard; protein; 110 AA.
 XX
 AC AAW15538;
 XX
 DT 27-NOV-1997 (first entry)
 XX
 DE Anti-TGF beta-1 scFv antibody 10A6 VL domain.
 XX
 KW Transforming growth factor beta-1; TGF-beta-1; human;

KW antibody engineering; scFv; phage display; lung fibrosis;
 KW arterial injury; proliferative retinopathy; retinal detachment;
 KW adult respiratory distress syndrome; liver cirrhosis;
 KW post myocardial infarction; post-angioplasty restenosis; scleroderma;
 KW vascular disease; cataract; glaucoma; scarring; glomerulonephritis;
 KW osteoporosis; immune disease; inflammation; rheumatoid arthritis;
 KW macrophage deficiency disease; macrophage pathogen infection; therapy.
 XX
 OS Homo sapiens.
 XX
 PN GB2305921-A.
 XX
 XX 23-APR-1997.
 PD
 XX
 XX 07-OCT-1996; 96GB-00020920.
 PF
 XX
 XX 06-OCT-1995; 95GB-00020486.
 PR
 XX 19-JAN-1996; 96GB-00001081.
 PR
 XX (CAME-) CAMBRIDGE ANTIBODY TECHNOLOGY.
 PA
 XX Thompson JE, Vaughan TJ, Williams AJ, Green JA, Jackson RH;
 PI Bacon L, Johnson KS, Wilton AJ, Tempest PR, Pope AR;
 PI
 XX WPI: 1997-215360/20.
 XX N-PSDB; AA160384.
 DR
 XX
 XX Agent contg. antigen-binding domain of human antibody to transforming
 PT growth factor beta 1 or 2 - and nucleic acid encoding it, used to
 PT neutralise effects of TGF, e.g. for control of fibrosis, immune and
 PT inflammatory disease.
 XX
 XX Claim 16; Fig 1b(ii); 184pp; English.
 PS
 XX This polypeptide sequence comprises the VL domain of human scFv antibody
 CC 10A6, which is specific for transforming growth factor (TGF) beta-1. It
 CC is encoded by a gene (AA160384) isolated from a peripheral blood
 CC lymphocyte library. The antigen-binding domains of human antibodies (see
 CC AA15522-40) to TGF beta-1 and/or beta-2 can be used to counter the
 CC adverse effects of TGF beta, such as (i) promotion of fibrosis (in
 CC dermal, ocular or keloid scarring, lung fibrosis, arterial injury,
 CC proliferative retinopathy, retinal detachment, adult respiratory distress
 CC syndrome, liver cirrhosis, post myocardial infarction, post-angioplasty
 CC restenosis, scleroderma, vascular disorders, cataract, glaucoma, or esp.
 CC neural scarring and glomerulonephritis, also (not claimed) osteoporosis,
 CC or (ii) immune and inflammatory diseases (e.g. rheumatoid arthritis,
 CC macrophage deficiency diseases or macrophage pathogen infection). Nucleic
 CC acids encoding human antibody VH and VL can be used for prodn. of
 CC recombinant antigen-binding domains. These are highly specific, have low
 CC dissociation constants (pref. less than 5 nM) and low IC50s for
 CC neutralisation
 XX
 SQ Sequence 110 AA;
 Query Match 93.0%; Score 53; DB 2; Length 110;
 Best Local Similarity 90.9%; Pred. No. 0.059;
 Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QGDSLKTYVAS 11
 Db 23 QGDSLKSYVAS 33
 RESULT 5
 ID ABP45519 standard; protein; 253 AA.
 XX
 XX AC ABP45519;
 XX
 XX 19-AUG-2002 (first entry)
 DT
 XX Human Blys binding scFv SEQ ID 1530.
 DE
 XX
 KW Blys; B lymphocyte stimulator; TNF superfamily; human; cytostatic;

KW Blys; B lymphocyte stimulator; TNF superfamily; human; cytostatic;
 KW tumour necrosis factor; B cell proliferation; B cell differentiation;
 KW immunosuppressive; immunostimulant; immunomodulatory; antirheumatic;
 KW antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;
 KW systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;
 KW common variable immunodeficiency; acquired immunodeficiency syndrome.
 XX
 OS Homo sapiens.
 XX
 PN WO200202641-A1.
 XX
 XX 10-JAN-2002.
 PD
 XX
 XX 15-JUN-2001; 2001WO-US019110.
 PF
 XX
 XX 16-JUN-2000; 2000US-0212210P.
 PR
 XX 17-OCT-2000; 2000US-0240816P.
 PR
 XX 16-MAR-2001; 2001US-0276248P.
 PR
 XX 21-MAR-2001; 2001US-0277379P.
 PR
 XX 25-MAY-2001; 2001US-0293499P.
 PR
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA (CAME-) CAMBRIDGE ANTIBODY TECHNOLOGY.
 PA
 XX Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;
 PI WPI: 2002-114799/15.
 PI
 XX
 XX Antibodies against B Lymphocyte Stimulating polypeptides, useful for the
 PT diagnosis and treatment of cancers and immune disorders.
 PT
 XX Claim 1; Page 2228-2229; 3148pp; English.
 PS
 XX This invention describes novel antibodies that immunospecifically bind to
 CC B Lymphocyte Stimulator (Blys) polypeptides. Blys is a member of the
 CC tumour necrosis factor (TNF) super family and induces B cell
 CC proliferation and differentiation. The antibodies of the invention have
 CC cytostatic, immunosuppressive, immunostimulant, immunomodulatory,
 CC antirheumatic and antiAIDS activity and can be used in vaccines to
 CC inhibit the expression and activity of Blys. The antibodies bind to Blys
 CC and so may be used to detect and quantitate the presence of Blys in
 CC biological samples and may be used in this way to diagnose disease
 CC associated with aberrant expression of Blys. They may also be
 CC administered to treat diseases associated with aberrant Blys expression
 CC and activity such as cancer, immune, and autoimmune disorders and
 CC diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis,
 CC immunodeficiency (e.g. common variable immunodeficiency (CVID) and
 CC acquired immunodeficiency syndrome (AIDS)). ABP43990-ABP47228 represent
 CC the antibodies and fragments of the antibodies described in the method of
 CC the invention
 XX
 SQ Sequence 253 AA;
 Query Match 93.0%; Score 53; DB 5; Length 253;
 Best Local Similarity 90.9%; Pred. No. 0.15;
 Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QGDSLKTYVAS 11
 Db 166 QGDSLKSYVAS 176
 RESULT 6
 ID ABP46073 standard; protein; 242 AA.
 XX
 XX AC ABP46073;
 XX
 XX 19-AUG-2002 (first entry)
 DT
 XX Human Blys binding scFv SEQ ID 2084.
 DE
 XX Blys; B lymphocyte stimulator; TNF superfamily; human; cytostatic;

KW tumour necrosis factor; B cell proliferation; B cell differentiation;
KW immunosuppressive; immunostimulant; immunomodulatory; antirheumatic;
KW antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;
KW systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;
KW common variable immunodeficiency; acquired immunodeficiency syndrome.
XX
OS Homo sapiens.
XX
PN WO200202641-A1.
XX
PD 10-JAN-2002.
XX
PF 15-JUN-2001; 2001WO-US019110.
XX
PR 16-JUN-2000; 2000US-0212210P.
PR 17-OCT-2000; 2000US-0240816P.
PR 16-MAR-2001; 2001US-0276248P.
PR 21-MAR-2001; 2001US-0277379P.
PR 25-MAY-2001; 2001US-0293499P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PA (CAMB-) CAMBRIDGE ANTIBODY TECHNOLOGY.
XX
PI Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;
XX
DR WPI; 2002-114799/15.
XX
PT Antibodies against B Lymphocyte Stimulating polypeptides, useful for the
PT diagnosis and treatment of cancers and immune disorders.
XX
PS Claim 1; Page 2883-2885; 3148pp; English.
XX
CC This invention describes novel antibodies that immunospecifically bind to
CC B Lymphocyte Stimulator (BlyS) polypeptides. BlyS is a member of the
CC tumour necrosis factor (TNF) super family and induces B cell
CC proliferation and differentiation. The antibodies of the invention have
CC cytostatic, immunosuppressive, immunostimulant, immunomodulatory,
CC antirheumatic and antiAIDS activity and can be used in vaccines to
CC inhibit the expression and activity of BlyS. The antibodies bind to BlyS
CC and so may be used to detect and quantitate the presence of BlyS in
CC biological samples and may be used in this way to diagnose disease
CC associated with aberrant expression of BlyS. They may also be
CC administered to treat diseases associated with aberrant BlyS expression
CC and activity such as cancer, immune, and autoimmune disorders and
CC diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis,
CC immunodeficiency (e.g. common variable immunodeficiency (CVID) and
CC acquired immunodeficiency syndrome (AIDS)). ABP43990-ABP47228 represent
CC the antibodies and fragments of the antibodies described in the method of
CC the invention
XX
SQ Sequence 242 AA;

Query Match 89.5%; Score 51; DB 5; Length 242;
Best Local Similarity 81.8%; Pred. No. 0.33;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
|||||:
DB 156 QGDSLRTYYAN 166

RESULT 7
ABP46091
ID ABP46091 standard; protein; 243 AA.
XX
AC ABP46091;
XX
DT 19-AUG-2002 (first entry)
XX
DE Human BlyS binding scFv SEQ ID 2102.

XX BlyS; B lymphocyte stimulator; TNF superfamily; human; cytostatic;
KW tumour necrosis factor; B cell proliferation; B cell differentiation;
KW immunosuppressive; immunostimulant; immunomodulatory; antirheumatic;

KW immunosuppressive; immunostimulant; immunomodulatory; antirheumatic;
KW antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;
KW systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;
KW common variable immunodeficiency; acquired immunodeficiency syndrome.
XX
OS Homo sapiens.
XX
PN WO200202641-A1.
XX
PD 10-JAN-2002.
XX
PF 15-JUN-2001; 2001WO-US019110.
XX
PR 16-JUN-2000; 2000US-0212210P.
PR 17-OCT-2000; 2000US-0240816P.
PR 16-MAR-2001; 2001US-0276248P.
PR 21-MAR-2001; 2001US-0277379P.
PR 25-MAY-2001; 2001US-0293499P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PA (CAMB-) CAMBRIDGE ANTIBODY TECHNOLOGY.
XX
PI Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;
XX
DR WPI; 2002-114799/15.
XX
PT Antibodies against B Lymphocyte Stimulating polypeptides, useful for the
PT diagnosis and treatment of cancers and immune disorders.
XX
PS Claim 1; Page 2905-2906; 3148pp; English.
XX
CC This invention describes novel antibodies that immunospecifically bind to
CC B Lymphocyte Stimulator (BlyS) polypeptides. BlyS is a member of the
CC tumour necrosis factor (TNF) super family and induces B cell
CC proliferation and differentiation. The antibodies of the invention have
CC cytostatic, immunosuppressive, immunostimulant, immunomodulatory,
CC antirheumatic and antiAIDS activity and can be used in vaccines to
CC inhibit the expression and activity of BlyS. The antibodies bind to BlyS
CC and so may be used to detect and quantitate the presence of BlyS in
CC biological samples and may be used in this way to diagnose disease
CC associated with aberrant expression of BlyS. They may also be
CC administered to treat diseases associated with aberrant BlyS expression
CC and activity such as cancer, immune, and autoimmune disorders and
CC diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis,
CC immunodeficiency (e.g. common variable immunodeficiency (CVID) and
CC acquired immunodeficiency syndrome (AIDS)). ABP43990-ABP47228 represent
CC the antibodies and fragments of the antibodies described in the method of
CC the invention
XX
SQ Sequence 243 AA;

Query Match 89.5%; Score 51; DB 5; Length 243;
Best Local Similarity 81.8%; Pred. No. 0.33;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
|||||:
DB 157 QGDSLRTYYAN 167

RESULT 8
ABP46092
ID ABP46092 standard; protein; 247 AA.
XX
AC ABP46092;
XX
DT 19-AUG-2002 (first entry)
XX
DE Human BlyS binding scFv SEQ ID 2103.

XX BlyS; B lymphocyte stimulator; TNF superfamily; human; cytostatic;
KW tumour necrosis factor; B cell proliferation; B cell differentiation;
KW immunosuppressive; immunostimulant; immunomodulatory; antirheumatic;

KW antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;
 KW systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;
 KW common variable immunodeficiency; acquired immunodeficiency syndrome.
 XX
 OS Homo sapiens.
 XX
 PN WO200202641-A1.
 XX
 PD 10-JAN-2002.
 XX
 PF 15-JUN-2001; 2001WO-US019110.
 XX
 PR 16-JUN-2000; 2000US-0212210P.
 XX
 PR 17-OCT-2000; 2000US-0240816P.
 PR 16-MAR-2001; 2001US-0276248P.
 PR 21-MAR-2001; 2001US-0277379P.
 PR 25-MAY-2001; 2001US-0293499P.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA (CAMB-) CAMBRIDGE ANTIBODY TECHNOLOGY.
 PA
 PI Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;
 XX WPI; 2002-114799/15.
 DR
 XX
 XX Antibodies against B Lymphocyte Stimulating polypeptides, useful for the
 PT diagnosis and treatment of cancers and immune disorders.
 PT
 XX
 XX Claim 1; Page 2906-2907; 3148pp; English.
 PS
 XX This invention describes novel antibodies that immunospecifically bind to
 CC B Lymphocyte Stimulator (BlyS) polypeptides. BlyS is a member of the
 CC tumour necrosis factor (TNF) super family and induces B cell
 CC proliferation and differentiation. The antibodies of the invention have
 CC cytostatic, immunosuppressive, immunostimulant, immunomodulatory,
 CC antirheumatic and antiAIDS activity and can be used in vaccines to
 CC inhibit the expression and activity of BlyS. The antibodies bind to BlyS
 CC and so may be used to detect and quantitate the presence of BlyS in
 CC biological samples and may be used in this way to diagnose disease
 CC associated with aberrant expression of BlyS. They may also be
 CC administered to treat diseases associated with aberrant BlyS expression
 CC and activity such as cancer, immune, and autoimmune disorders and
 CC diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis,
 CC immunodeficiency (e.g. common variable immunodeficiency (CVID) and
 CC acquired immunodeficiency syndrome (AIDS)). ABP43990-ABP47228 represent
 CC the antibodies and fragments of the antibodies described in the method of
 CC the invention
 XX
 XX Sequence 247 AA;
 SQ
 Query Match 89.5%; Score 51; DB 5; Length 247;
 Best Local Similarity 81.8%; Pred. No. 0.33;
 Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QGDSLKTYVAS 11
 DB 161 QGDSLRTYYAN 171
 RESULT 9
 ABP46077
 ID ABP46077 standard; protein; 248 AA.
 XX
 AC ABP46077;
 XX
 DT 19-AUG-2002 (first entry)
 XX
 DE Human BlyS binding scFv SEQ ID 2088.
 XX
 XX BlyS; B lymphocyte stimulator; TNF superfamily; human; cytostatic;
 KW tumour necrosis factor; B cell proliferation; B cell differentiation;
 KW immunosuppressive; immunostimulant; immunomodulatory; antirheumatic;
 KW antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;
 KW

KW systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;
 KW common variable immunodeficiency; acquired immunodeficiency syndrome.
 XX
 OS Homo sapiens.
 XX
 PN WO200202641-A1.
 XX
 PD 10-JAN-2002.
 XX
 PF 15-JUN-2001; 2001WO-US019110.
 XX
 PR 16-JUN-2000; 2000US-0212210P.
 XX
 PR 17-OCT-2000; 2000US-0240816P.
 PR 16-MAR-2001; 2001US-0276248P.
 PR 21-MAR-2001; 2001US-0277379P.
 PR 25-MAY-2001; 2001US-0293499P.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA (CAMB-) CAMBRIDGE ANTIBODY TECHNOLOGY.
 PA
 PI Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;
 XX WPI; 2002-114799/15.
 DR
 XX
 XX Antibodies against B Lymphocyte Stimulating polypeptides, useful for the
 PT diagnosis and treatment of cancers and immune disorders.
 PT
 XX
 XX Claim 1; Page 2888-2889; 3148pp; English.
 PS
 XX This invention describes novel antibodies that immunospecifically bind to
 CC B Lymphocyte Stimulator (BlyS) polypeptides. BlyS is a member of the
 CC tumour necrosis factor (TNF) super family and induces B cell
 CC proliferation and differentiation. The antibodies of the invention have
 CC cytostatic, immunosuppressive, immunostimulant, immunomodulatory,
 CC antirheumatic and antiAIDS activity and can be used in vaccines to
 CC inhibit the expression and activity of BlyS. The antibodies bind to BlyS
 CC and so may be used to detect and quantitate the presence of BlyS in
 CC biological samples and may be used in this way to diagnose disease
 CC associated with aberrant expression of BlyS. They may also be
 CC administered to treat diseases associated with aberrant BlyS expression
 CC and activity such as cancer, immune, and autoimmune disorders and
 CC diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis,
 CC immunodeficiency (e.g. common variable immunodeficiency (CVID) and
 CC acquired immunodeficiency syndrome (AIDS)). ABP43990-ABP47228 represent
 CC the antibodies and fragments of the antibodies described in the method of
 CC the invention
 XX
 XX Sequence 248 AA;
 SQ
 Query Match 89.5%; Score 51; DB 5; Length 248;
 Best Local Similarity 81.8%; Pred. No. 0.33;
 Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QGDSLKTYVAS 11
 DB 162 QGDSLRTYYAN 172
 RESULT 10
 ABP45991
 ID ABP45991 standard; protein; 249 AA.
 XX
 AC ABP45991;
 XX
 DT 19-AUG-2002 (first entry)
 XX
 DE Human BlyS binding scFv SEQ ID 2002.
 XX
 XX BlyS; B lymphocyte stimulator; TNF superfamily; human; cytostatic;
 KW tumour necrosis factor; B cell proliferation; B cell differentiation;
 KW immunosuppressive; immunostimulant; immunomodulatory; antirheumatic;
 KW antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;
 KW systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;

common variable immunodeficiency; acquired immunodeficiency syndrome.

XX Homo sapiens.
 OS WO200202641-A1.
 PN 10-JAN-2002.
 XX 15-JUN-2001; 2001WO-US019110.
 XX 16-JUN-2000; 2000US-0212210P.
 XX 17-OCT-2000; 2000US-0240816P.
 PR 16-MAR-2001; 2001US-0276248P.
 PR 21-MAR-2001; 2001US-0277379P.
 PR 25-MAY-2001; 2001US-0293499P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA (CAMP-) CAMBRIDGE ANTIBODY TECHNOLOGY.
 XX Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;
 PI WPI; 2002-114799/15.
 XX Antibodies against B Lymphocyte Stimulating polypeptides, useful for the
 PT diagnosis and treatment of cancers and immune disorders.
 XX Claim 1; Page 2789-2790; 3148pp; English.
 XX This invention describes novel antibodies that immunospecifically bind to
 CC B Lymphocyte Stimulator (Blys) polypeptides. Blys is a member of the
 CC tumour necrosis factor (TNF) super family and induces B cell
 CC proliferation and differentiation. The antibodies of the invention have
 CC cytostatic, immunosuppressive, immunostimulant, immunomodulatory,
 CC antirheumatic and antiAIDS activity and can be used in vaccines to
 CC inhibit the expression and activity of Blys. The antibodies bind to Blys
 CC and so may be used to detect and quantitate the presence of Blys in
 CC biological samples and may be used in this way to diagnose disease
 CC associated with aberrant expression of Blys. They may also be
 CC administered to treat diseases associated with aberrant Blys expression
 CC and activity such as cancer, immune, and autoimmune disorders and
 CC diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis,
 CC immunodeficiency (e.g. common variable immunodeficiency (CVID) and
 CC acquired immunodeficiency syndrome (AIDS)). ABP43990-ABP47228 represent
 CC the antibodies and fragments of the antibodies described in the method of
 CC the invention
 XX Sequence 249 AA;
 SQ Query Match 89.5%; Score 51; DB 5; Length 249;
 Best Local Similarity 81.8%; Pred. No. 0.34;
 Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QGDSLKTYVAS 11
 DB 163 QGDSLRTYYAN 173
 RESULT 11
 ABP45979
 ID ABP45979 standard; protein; 250 AA.
 XX AC ABP45979;
 XX 19-AUG-2002 (first entry)
 DT Human Blys binding scFv SEQ ID 1990.
 XX Blys; B lymphocyte stimulator; TNF superfamily; human; cytostatic;
 XX tumour necrosis factor; B cell proliferation; B cell differentiation;
 XX immunosuppressive; immunostimulant; immunomodulatory; antirheumatic;
 XX antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;
 XX systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;
 KW common variable immunodeficiency; acquired immunodeficiency syndrome.

XX Homo sapiens.
 OS WO200202641-A1.
 PN 10-JAN-2002.
 XX 15-JUN-2001; 2001WO-US019110.
 XX 16-JUN-2000; 2000US-0212210P.
 XX 17-OCT-2000; 2000US-0240816P.
 PR 16-MAR-2001; 2001US-0276248P.
 PR 21-MAR-2001; 2001US-0277379P.
 PR 25-MAY-2001; 2001US-0293499P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA (CAMP-) CAMBRIDGE ANTIBODY TECHNOLOGY.
 XX Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;
 PI WPI; 2002-114799/15.
 XX Antibodies against B Lymphocyte Stimulating polypeptides, useful for the
 PT diagnosis and treatment of cancers and immune disorders.
 XX Claim 1; Page 2774-2775; 3148pp; English.
 XX This invention describes novel antibodies that immunospecifically bind to
 CC B Lymphocyte Stimulator (Blys) polypeptides. Blys is a member of the
 CC tumour necrosis factor (TNF) super family and induces B cell
 CC proliferation and differentiation. The antibodies of the invention have
 CC cytostatic, immunosuppressive, immunostimulant, immunomodulatory,
 CC antirheumatic and antiAIDS activity and can be used in vaccines to
 CC inhibit the expression and activity of Blys. The antibodies bind to Blys
 CC and so may be used to detect and quantitate the presence of Blys in
 CC biological samples and may be used in this way to diagnose disease
 CC associated with aberrant expression of Blys. They may also be
 CC administered to treat diseases associated with aberrant Blys expression
 CC and activity such as cancer, immune, and autoimmune disorders and
 CC diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis,
 CC immunodeficiency (e.g. common variable immunodeficiency (CVID) and
 CC acquired immunodeficiency syndrome (AIDS)). ABP43990-ABP47228 represent
 CC the antibodies and fragments of the antibodies described in the method of
 CC the invention
 XX Sequence 250 AA;
 SQ Query Match 89.5%; Score 51; DB 5; Length 250;
 Best Local Similarity 81.8%; Pred. No. 0.34;
 Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QGDSLKTYVAS 11
 DB 164 QGDSLRTYYAN 174
 RESULT 12
 AAY95195
 ID AAY95195 standard; peptide; 11 AA.
 XX AC AAY95195;
 XX 29-AUG-2000 (first entry)
 DT Anti-platelet glycoprotein Ib human H1b-1 VL CDR1.
 XX Variable light chain; single chain antibody; scFv; human; H1b-1;
 KW glycoprotein Ib alpha; platelet; aggregation; antiaggregant;
 KW antithrombotic; thrombus; therapy; diagnostic; CDR1;
 XX complementarity determining region.
 OS Homo sapiens.
 XX

PN WO200026667-A1.
 XX 11-MAY-2000.
 XX 29-OCT-1999; 99WO-US025495.
 XX 30-OCT-1998; 98US-0106275P.
 XX (MILL/) MILLER J L.
 XX Miller JL;
 XX WPI; 2000-365744/31.
 XX Isolated nucleic acid molecule encoding anti-human platelet glycoprotein
 PT Ib alpha molecule useful for producing antibodies which inhibit platelet
 PT aggregation.
 XX Claim 19; Fig 5; 89pp; English.
 XX The present sequence is that of complementarity determining region 1
 CC (CDR1) of the light chain variable region (VL) of human single chain
 CC antibody (scFv) Hib-1 (see AAY95198), which is directed against platelet
 CC glycoprotein Ib (GPIb). The Hib series of scFv was isolated from a human
 CC synthetic VH and VL scFv library on the basis of their binding to
 CC platelet GPIb. Whether displayed as surface proteins on a phagemid or
 CC secreted as free scFv by Escherichia coli, the Hib scFv clones are
 CC capable of inhibiting von Willebrand factor-dependent aggregation of
 CC platelets. The scFv are composed of native human protein sequences and
 CC are therefore attractive potential reagents for therapeutic purposes.
 CC They provide a new class of antithrombotic agents, useful for the
 CC prevention of platelet-dependent thrombi in diseased arteries, bypass
 CC grafts, dialysis etc., and can also be used as diagnostic reagents.
 CC Methods of inhibiting aggregation of platelets, of binding human platelet
 CC GPIb alpha and of selecting a VH or VL region of an antibody that
 CC inhibits platelet aggregation are claimed. Fragments of the scFv VH or VL
 CC chain, including CDR fragments, are also claimed.
 XX Sequence 11 AA;
 SQ
 Query Match 87.7%; Score 50; DB 3; Length 11;
 Best Local Similarity 81.8%; Pred. No. 0.018;
 Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QGDSLKTYVAS 11
 DB |||||:||||
 1 QGDSLRSYYAS 11
 RESULT 13
 ID AAY95216
 XX AAY95216 standard; peptide; 11 AA.
 AC AAY95216;
 XX 29-AUG-2000 (first entry)
 XX Anti-platelet glycoprotein Ib human Hib-3 VL CDR1.
 DE Variable light chain; single chain antibody; scFv; human; Hib-3;
 XX glycoprotein Ib alpha; platelet; aggregation; antiaggregant;
 KW antithrombotic; thrombus; therapy; diagnostic; CDR1;
 KW complementarity determining region.
 XX Homo sapiens.
 OS WO200026667-A1.
 XX 11-MAY-2000.
 XX 29-OCT-1999; 99WO-US025495.
 XX 30-OCT-1998; 98US-0106275P.

XX (MILL/) MILLER J L.
 XX Miller JL;
 XX WPI; 2000-365744/31.
 XX Isolated nucleic acid molecule encoding anti-human platelet glycoprotein
 PT Ib alpha molecule useful for producing antibodies which inhibit platelet
 PT aggregation.
 XX Claim 19; Fig 7; 89pp; English.
 XX The present sequence is that of complementarity determining region 1
 CC (CDR1) of the light chain variable region (VL) of human single chain
 CC antibody (scFv) Hib-3 (see AAY95219), which is directed against platelet
 CC glycoprotein Ib (GPIb). The Hib series of scFv was isolated from a human
 CC synthetic VH and VL scFv library on the basis of their binding to
 CC platelet GPIb. Whether displayed as surface proteins on a phagemid or
 CC secreted as free scFv by Escherichia coli, the Hib scFv clones are
 CC capable of inhibiting von Willebrand factor-dependent aggregation of
 CC platelets. The scFv are composed of native human protein sequences and
 CC are therefore attractive potential reagents for therapeutic purposes.
 CC They provide a new class of antithrombotic agents, useful for the
 CC prevention of platelet-dependent thrombi in diseased arteries, bypass
 CC grafts, dialysis etc., and can also be used as diagnostic reagents.
 CC Methods of inhibiting aggregation of platelets, of binding human platelet
 CC GPIb alpha and of selecting a VH or VL region of an antibody that
 CC inhibits platelet aggregation are claimed. Fragments of the scFv VH or VL
 CC chain, including CDR fragments, are also claimed.
 XX Sequence 11 AA;
 SQ
 Query Match 87.7%; Score 50; DB 3; Length 11;
 Best Local Similarity 81.8%; Pred. No. 0.018;
 Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QGDSLKTYVAS 11
 DB |||||:||||
 1 QGDSLRSYYAS 11
 RESULT 14
 ID AAG80192
 XX AAG80192 standard; peptide; 11 AA.
 AC AAG80192;
 XX 22-JAN-2002 (first entry)
 XX Human gp96 antibody variable light chain 'CDR1' peptide.
 DE Antibody; light chain; variable region; CDR1; human; gp96; cytostatic;
 KW complementarity determining region; antibacterial; virucide; fungicide;
 KW protozoacide; tumour; autologous immunization; anti-idiotypic.
 XX Homo sapiens.
 OS DE10019967-A1.
 XX 25-OCT-2001.
 XX 20-APR-2000; 2000DE-01019967.
 XX 20-APR-2000; 2000DE-01019967.
 XX (FENN-) FENNING BIOMED GMBH.
 XX Terness P, Kleist C, Opelz G, Welschhof M, Arnold-Schild D;
 PI Schild H, Rammensee H;
 XX WPI; 2002-018649/03.
 XX

PT New recombinant antibody specific for native gp96, useful for isolating
PT gp96 as complex with endogenous peptides, used for autologous
PT immunotherapy of e.g. tumor.
XX
XX
XX Claim 3; Page 13; 28pp; German.
XX
XX This invention describes a novel recombinant antibody (Ab) that binds
CC native gp96. The product of the invention has cytostatic, antibacterial,
CC virucide, fungicide and protozoacide activity. Complexes of gp96 with the
CC peptides of the invention induce an immune response (cytotoxic T cells)
CC specific for the associated peptides. Ab are used for purification or
CC labeling of gp96, including intact complexes (C) with peptides (I), from
CC small amounts of tumour or infected cells. (C) are useful: (i) in human
CC or veterinary medicine for autologous immunization, for treating tumours
CC and infections (by viruses, bacteria, mycoplasma, fungi and parasites
CC (protozoa)); and (ii) for raising anti-idiotypic antibodies. Ab make
CC possible isolation of very pure gp96 in native form, even from small
CC tumour samples, in a one-step chromatographic process that does not
CC require concanavalin A (as used in known processes and which may
CC contaminate the final product by bleeding from the column). This sequence
CC represents the variable light chain complementarity determining region
CC (CDR1') fragment of the antibody raised against human gp96 described in
CC the method of the invention
XX
XX Sequence 11 AA;
SQ

Query Match 87.7%; Score 50; DB 5; Length 11;
Best Local Similarity 81.8%; Pred. No. 0.018;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
DB 1 QGDSLRSYVAS 11

RESULT 15
AAR22572
ID AAR22572 standard; protein; 101 AA.
XX
XX AAR22572;
XX
XX 21-MAY-1992 (first entry)

XX Light chain VL3.5 from BSA binding scfv fragment.
XX
XX Fd; bacteriophage; gene III; filamentous; phagemid; capsid; coat; pilus;
KW 33p; binding; adsorption; gene VIII; diverse repertoire;
KW specific binding pairs; replicable genetic display package; human.
XX
XX Homo sapiens.
XX
XX WO9201047-A.
XX
XX 23-JAN-1992.
XX
XX 10-JUL-1990; 90GB-00015198.
XX
XX 10-JUL-1990; 90GB-00015198.
PR 19-OCT-1990; 90GB-00022845.
PR 12-NOV-1990; 90GB-00024503.
PR 06-MAR-1991; 91GB-00004744.
PR 15-MAY-1991; 91GB-00010549.
XX
XX (CAMP-) CAMBRIDGE ANTIBODY TECHNOLOGY.
PA (MEDI-) MED RES COUNCIL.

XX McCafferty J, Pope AR, Johnson KS, Hoogenboom HRJ, Griffiths AD;
PI Jackson RH, Holliger KP, Marks JD;
XX WPI; 1992-056862/07.
XX
XX Producing members of specific binding pairs - by expression in
PT recombinant host cells with a secreting replicable genetic display

PT package.
XX
XX Table 11; Page 152; 209pp; English.

XX PCR was used to prepare a human scFv library from RNA from white blood
CC cells from an unimmunised donor. Heavy chains from IgG and IgM antibodies
CC were amplified separately. Four separate libraries were generated (IgG-K,
CC IgG-lambda, IgM-K and IgM-lambda). The purified scFv fragments were
CC ligated into the phagemid pHE1 for expression on the surface of fd
CC bacteriophage as gene III fusions. The clones were then subjected to
CC affinity selection for binding to phOx:BSA by selection on tubes followed
CC by analysis by ELISA. Of 96 clones analysed, 43 showed binding to both
CC phOx:BSA and BSA. These were designated BSA binders. Thirteen of fourteen
CC clones sequenced had the same sequence, the VH derived from a human VH3
CC family gene (AAR22571) and the VL from a human V lambda 3 family gene
CC (shown here). The other was derived from a human VH4 family gene and a
CC human VK1 family gene. One clone bound only to phOx:BSA (oxazolone
CC binder). This sequence revealed a VH derived from a human VH1 family gene
CC (AAR22569) and VL from a human V lambda 1 family gene (AAR22570). See
CC also AAR21260-307, 309-312, AAR22450, AAR22565, AAR22567-81
XX
XX Sequence 101 AA;
SQ

Query Match 87.7%; Score 50; DB 2; Length 101;
Best Local Similarity 81.8%; Pred. No. 0.19;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
DB 23 QGDSLRSYVAS 33

RESULT 16
AAR80091
ID AAR80091 standard; protein; 103 AA.
XX
XX AAR80091;
XX
XX 23-MAY-1996 (first entry)

XX Human derived light chain RT3 phage antibody.
XX
XX Light chain; RT3; human; catalytic antibody; bacteriophage.
XX
XX Homo sapiens.
XX
XX Key.

XX Location/Qualifiers
FT Region 1..16
FT /note= "framework region 1"
FT Region 17..27
FT /note= "complementarity determining region 1"
FT Region 28..43
FT /note= "framework region 2"
FT Region 44..49
FT /note= "complementarity determining region 2"
FT Region 50..82
FT /note= "framework region 3"
FT Region 83..92
FT /note= "complementarity determining region 3"
FT Region 93..103
FT /note= "framework region 4"
XX
XX WO9527045-A1.
XX
XX 12-OCT-1995.
XX
XX 30-MAR-1994; 94WO-US003420.
XX
XX 30-MAR-1994; 94WO-US003420.
XX
XX (IGEN-) IGEN INC.
XX
XX Smith RG, McCafferty J, Chiswell D, Darsley MJ, Fitzgerald K;

PI Kenten JH, Martin MT, Titmas RC, Williams RO;
 XX WPI: 1995-358624/46.
 DR N-PSDB; AAT04638.
 XX
 PT Production of catalytic antibodies displayed on phage - by generating a
 PT gene library of antibody-derived domains and expressing it in phage
 PT vectors.
 XX
 PS Disclosure; Fig 20; 133pp; English.
 XX
 CC AAT04638 encodes AAR80091 human derived light chain RT3 phage antibody.
 CC The DNA was used in the prepn. of catalytic antibody (CA) producing
 CC bacteriophage. The CAs can be used to activate/deactivate a biological
 CC function in an animal by enhancing the rate of cleavage, or formation of
 CC a specific bond within a mol. in vivo
 XX
 SQ Sequence 103 AA;
 Query Match 87.7%; Score 50; DB 2; Length 103;
 Best Local Similarity 81.8%; Pred. No. 0.2;
 Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QGDSLXTYYAS 11
 |||||:||||
 Db 17 QGDSLRSYYAS 27
 RESULT 17
 AAW95489
 ID AAW95489 standard; protein; 103 AA.
 XX
 AC AAW95489;
 DT 29-MAR-1999 (first entry)
 XX
 DE Human-derived RT3 phage antibody light chain genetic sequence.
 XX
 XX Catalytic; antibody; phage display; immunising; phage expression vector;
 KW prodrug; scFV; RT3.
 XX
 OS Homo sapiens.
 XX
 PN US855885-A.
 XX
 PD 05-JAN-1999.
 XX
 PF 14-JUL-1994; 94US-00273146.
 XX
 PR 22-JAN-1993; 93US-00007684.
 XX
 PA (MCCA/) MCCAFFERTY J.
 PA (CHIS/) CHISWELL D.
 PA (DARS/) DARSLEY M J.
 PA (TITM/) TITMAS R C.
 PA (NART/) MARTIN M T.
 PA (KENT/) KENTEN J H.
 PA (SMIT/) SMITH R.
 PA (FITZ/) FITZGERALD K.
 PA (WILL/) WILLIAMS R O.
 XX
 XX Fitzgerald K, Darsley MJ, Williams RO, Smith R, Martin MT;
 PI Kenten JH, Chiswell D, McCafferty J, Titmas RC;
 XX WPI: 1999-105036/09.
 DR N-PSDB; AAX00888.
 XX
 XX Production of catalytic antibodies displayed on bacteriophages -
 PT comprises generating a gene library of antibody-derived domains inserting
 PT coding into a phage expression vector and isolating the catalytic
 PT antibodies.
 XX
 PS Example; Fig 20F; 117pp; English.

XX The invention relates to methods for producing catalytic antibodies
 CC displayed on a phage. The method comprises: (a) generating a gene library
 CC of antibody-derived domains; (b) inserting coding for the domains into a
 CC phage expression vector; and (c) isolating the catalytic antibodies. The
 CC phage expression vector incorporates a histidine peptide in tandem with a
 CC myc peptide. The catalytic antibodies can be isolated by preparing an
 CC antigen; optionally immunising an animal with the antigen; generating a
 CC library of VH and VL domains from the immunised animal; cloning the VH
 CC and VL domains into a phage expression vector to generate phage display
 CC antibodies; selecting phage display antibodies which bind specifically to
 CC the antigen; screening the selected phage display antibodies for
 CC catalytic activity to substrate; and isolating the catalytic antibodies,
 CC where the phage expression vector incorporates a histidine peptide in
 CC tandem with a myc peptide. The processes are used to produce catalytic
 CC antibodies, which can be used for in vivo activation of a prodrug.
 CC Sequences AAW95484-489 represent genetic sequences of heavy and light
 CC chains of RT3 specific phage antibodies selected from a naive human phage
 CC antibody library
 XX
 SQ Sequence 103 AA;
 Query Match 87.7%; Score 50; DB 2; Length 103;
 Best Local Similarity 81.8%; Pred. No. 0.2;
 Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QGDSLXTYYAS 11
 |||||:||||
 Db 17 QGDSLRSYYAS 27
 RESULT 18
 AAG93606
 ID AAG93606 standard; protein; 104 AA.
 XX
 AC AAG93606;
 XX
 DT 14-SEP-2001 (first entry)
 XX
 DE Human anti-Rh(D) chain J04 protein sequence.
 XX
 KW Human; anti-Rh(D) antibody; Rh(D) binding protein; immunostimulant;
 KW red blood cell; Rh phenotype; diagnosis; therapeutic.
 XX
 OS Homo sapiens.
 XX
 PN US6255455-B1.
 XX
 PD 03-JUL-2001.
 XX
 PF 29-JAN-1999; 99US-00240274.
 XX
 PR 11-OCT-1996; 96US-0028550P.
 PR 27-JUN-1997; 97US-00884045.
 PR 10-APR-1998; 98US-0081380P.
 XX
 PA (UYPE-) UNIV PENNSYLVANIA.
 XX
 PI Siegel DL;
 XX
 DR WPI: 2001-388931/41.
 DR N-PSDB; AAH68663.
 XX
 PT New isolated protein, preferably a human anti-Rh(D) antibody for use in
 PT diagnostics requiring a human instead of an animal antibody and in
 PT therapeutic medicine.
 XX
 PS Claim 1; Col 45; 162pp; English.
 XX
 CC The present invention describes an isolated Rh(D) binding protein,
 CC preferably a human antibody, (i) having an amino acid sequence comprising
 CC one of the sequences (S) given in AAG93558 to AAG93669..(i) has
 CC immunostimulant activity, and can be used as an immune system stimulant.

CC (I) can be used in diagnostic and therapeutic medicine. The antibodies
 CC are used in diagnostics that require human antibodies instead of animal
 CC antibodies, such as determine the Rh phenotype of human red blood cells.
 CC AAH68615 to AAH68726 represent the nucleotide sequence which encode
 CC AAG93558 to AAG93669. AAG93670 to AAG93697 represent anti-Rh(D) heavy
 CC chain CDR3 amino acid sequences which are given in the exemplification of
 CC the present invention
 XX
 SQ Sequence 104 AA;

Query Match 87.7%; Score 50; DB 4; Length 104;
 Best Local Similarity 81.8%; Pred. No. 0.2;
 Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
 |||||:||||
 Db 21 QGDSLRSYVAS 31

RESULT 19
 ABB05054
 ID ABB05054 standard; protein; 104 AA.

XX AC ABB05054;
 XX DT 26-MAR-2002 (first entry)
 XX DE Fibronectin isoform B specific binding antibody VL domain protein.
 XX KW Fibronectin isoform B; B-FN; extra domain B; ED-B; antibody; VL domain;
 XX KW angiogenesis; neoplasia; tumour; detection.

XX OS Unidentified.
 XX PN WO200196599-A2.

XX PD 20-DEC-2001.

XX PF 08-JUN-2001; 2001WO-BP006533.

XX PR 15-JUN-2000; 2000US-0211677P.

XX PA (PHIL-) PHILOGEN SRL.

XX PA (CAST/) CASTELLANI P.

XX PA (ZARD/) ZARDI L.

XX PA (ZIJL/) ZIJLSTRA A.

XX PI Castellani P, Zardi L, Zijlstra A;

XX WPI; 2002-098076/13.

XX DR N-PSDB; ABA92718.

XX PT Detecting fibronectin isoform B in body fluid of an individual, by
 PT binding fibronectin to gelatin on solid support, and determining ability
 PT of specific binding member for Extra Domain B to be retained on support.

XX PS Disclosure; Page 56; 72pp; English.

XX CC The present invention describes a method for detecting the presence of
 CC fibronectin isoform B (B-FN) in body fluid of an individual. The method
 CC involves passing the fluid over a solid support (SS) to which gelatin
 CC that binds FN is attached, such that if FN is present, it is retained on
 CC SS by binding gelatin, and determining the presence of B-FN on SS by
 CC determining the ability of a specific binding member for extra domain B
 CC (ED-B) of B-FN to be retained on SS. The method is useful for the
 CC quantitative determination of B-FN in a body fluid taken or collected
 CC from an individual. Other methods from the present invention can be used:
 CC for quantitating the amount of an insoluble marker protein in a sample of
 CC tissue or tumour; for determining the grade or severity of neoplasia in a
 CC sample of tissue or tumour. An elevated level of B-FN in a body fluid
 CC such as plasma, cerebral-spinal fluid or cystic fluid obtained from an
 CC individual is indicative of the presence of a tumour in the individual.
 CC The methods are useful in diagnostic and prognostic situations. The

CC methods can be readily applied to small biopsy samples obtained in the
 CC clinic, thereby providing valuable information without additional
 CC compromise to the patients or complications for the clinicians. The
 CC present sequence represents the VL domain of a B-FN ED-B specific binding
 CC antibody molecule, which is used in the exemplification of the present
 CC invention
 XX
 SQ Sequence 104 AA;

Query Match 87.7%; Score 50; DB 5; Length 104;
 Best Local Similarity 81.8%; Pred. No. 0.2;
 Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
 |||||:||||
 Db 23 QGDSLRSYVAS 33

RESULT 20
 ABO27413
 ID ABO27413 standard; protein; 104 AA.

XX AC ABO27413;

XX DT 12-SEP-2003 (first entry)

XX DE Anti-Rh(D) chain J04.

XX KW Human; RH(D) binding protein; blood typing; blood product; antibody;
 XX KW magnetically activated cell sorting.

XX OS Homo sapiens.

XX PN US2003040605-A1.

XX PD 27-FEB-2003.

XX PF 04-MAY-2001; 2001US-00848798.

XX PR 11-OCT-1996; 96US-0028550P.

XX PR 27-JUN-1997; 97US-00884045.

XX PR 10-APR-1998; 98US-0081380P.

XX PR 29-JAN-1999; 99US-00240274.

XX PA (UYPE-) UNIV PENNSYLVANIA.

XX PI Siegel DL;

XX WPI; 2003-512273/48.

XX DR N-PSDB; ACD45327.

XX PT New human Rh(D)-binding protein useful for various diagnostic and
 XX therapeutic applications, including typing of blood or blood products.

XX PS Claim 4; Page 28; 187pp; English.

XX CC The invention relates to an isolated Rh(D) binding protein. The protein
 CC can be used for magnetically activated cell sorting. The protein is
 CC useful in various diagnostic and therapeutic applications in humans,
 CC including typing of blood or blood products. The present sequence
 CC represents the amino acid sequence of a human anti-Rh(D) chain

XX SQ Sequence 104 AA;

Query Match 87.7%; Score 50; DB 6; Length 104;
 Best Local Similarity 81.8%; Pred. No. 0.2;
 Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
 |||||:||||
 Db 21 QGDSLRSYVAS 31

RESULT 21

AAU02531
 ID AAU02531 standard; protein; 106 AA.
 XX AC AAU02531;
 XX DT 29-AUG-2001 (first entry)
 XX DE Anti-adipocyte monoclonal antibody light chain, FAT 20.
 XX KW Antibody; adipocyte; heavy chain; light chain; obesity; fat;
 XX KW heart disease; complementarity determining region; CDR.
 XX OS Homo sapiens.
 XX PN WO200127279-A1.
 XX PD 19-APR-2001.
 XX PF 11-OCT-2000; 2000WO-GB003900.
 XX PR 12-OCT-1999; 99US-0158812P.
 XX PA (CAME-) CAMBRIDGE ANTIBODY TECHNOLOGY.
 XX PI Edwards BM, Main SH, Vaughan TJ;
 XX DR WPI; 2001-282031/29.
 XX DR N-PSDB; AAS03431.
 XX PT Panel of specific binding members of antibody molecules which bind to
 XX PT whole adipocytes is used in the treatment of obesity and obesity related
 XX PT diseases.
 XX PS Claim 1; Page 112; 182pp; English.

Claim 1; Page 112; 182pp; English.

AAU02501-AAU02635, and AAU02641-AAU02748 represent the amino acid
 sequences of anti-adipocyte monoclonal antibody heavy chain, light chain,
 and heavy chain complementarity determining regions (CDR) of the
 invention. The antibodies can be used in the treatment of obesity and
 obesity related diseases. The antibodies can be used to deliver drugs or
 pro-drugs directly to the fat mass of an obese patient or the antibody
 can be used as a therapeutic itself. Antibodies binding specifically to
 adipocytes can be used to activate the immune system to destroy the cells
 by complement mediated lysis. The antibodies may be labeled with a
 detectable label such as radiolabel, fluorescent or chemical group and
 used in methods of diagnosis in human subjects e.g. to determine the
 presence of adipocyte antigen on the surface of an adipocyte to detect or
 determine the presence or level of adipocytes in a cell or tissue sample.
 The antibodies can be used as an alternative means of treatment for obese
 patients other than undergoing surgery to remove excess fat. Antibodies
 for different types of fat deposits can also be produced e.g. intra-
 abdominal fat associated with heart disease

Sequence 106 AA;

Query Match 87.7%; Score 50; DB 4; Length 106;
 Best Local Similarity 81.8%; Pred. No. 0.2;
 Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11

DB 23 QGDSLSRYAS 33

RESULT 22

AAV95215
 ID AAV95215 standard; protein; 107 AA.
 XX AC AAV95215;
 XX DT 29-AUG-2000 (first entry)

QY 1 QGDSLKTYVAS 11

DB 23 QGDSLSRYAS 33

DE Anti-platelet glycoprotein Ib human Hib-3 VL.

XX KW Variable light chain; single chain antibody; scFv; human; Hib-3;
 KW glycoprotein Ib alpha; platelet; aggregation; antiaggregant;
 KW antithrombotic; thrombus; therapy; diagnostic.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Region 1. .22
 FT /note= "framework region 1"
 FT Region 23. .33
 FT /note= "complementarity determining region 1"
 FT Region 34. .48
 FT /note= "framework region 2"
 FT Region 49. .55
 FT /note= "complementarity determining region 2"
 FT Region 56. .87
 FT /note= "framework region 3"
 FT Region 88. .96
 FT /note= "complementarity determining region 3"
 FT Region 97. .107
 FT /note= "framework region 4"

WO200026667-A1.

11-MAY-2000.

29-OCT-1999; 99WO-US025495.

30-OCT-1998; 98US-0106275P.

(MILL/) MILLER J L.

Miller JL;

WPI; 2000-365744/31.

Isolated nucleic acid molecule encoding anti-human platelet glycoprotein
 Ib alpha molecule useful for producing antibodies which inhibit platelet
 aggregation.

Claim 18; Fig 7; 89pp; English.

The present sequence is that of the light chain variable region (VL) of
 human single chain antibody (scFv) Hib-3 (see AAY95219), which is
 directed against platelet glycoprotein Ib (GPIb). The Hib series of scFv
 was isolated from a human synthetic VH and VL scFv library by 3 rounds of
 phagemid selection against transfected CHO cells expressing the GPIb
 alpha component of the GPIb/IX/V complex on their surface, followed by a
 4th round of selection against washed human platelets, and 2 final rounds
 in which attempts were made to displace scFv from washed platelets by
 flooding with murine monoclonal antibody or mimotope peptide (see
 AAY95229). Whether displayed as surface proteins on a phagemid or
 secreted as free scFv by *Escherichia coli*, the Hib scFv clones are
 capable of inhibiting von Willebrand factor-dependent aggregation of
 platelets. The scFv are composed of native human protein sequences and
 are therefore attractive potential reagents for therapeutic purposes.
 They provide a new class of antithrombotic agents, useful for the
 prevention of platelet-dependent thrombi in diseased arteries, bypass
 grafts, dialysis etc., and can also be used as diagnostic reagents.
 Methods of inhibiting aggregation of platelets, of binding human platelet
 GPIb alpha and of selecting a VH or VL region of an antibody that
 inhibits platelet aggregation are claimed

Sequence 107 AA;

Query Match 87.7%; Score 50; DB 3; Length 107;
 Best Local Similarity 81.8%; Pred. No. 0.21;
 Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11

|||||::|||

Db 23 QGDSLSRYSYAS 33

RESULT 23
AAV95194

ID AAY95194 standard; peptide; 107 AA.
XX AC AAY95194;
XX 29-AUG-2000 (first entry)
XX DE Anti-platelet glycoprotein Ib human H1b-1 VL.
XX DE Variable light chain; single chain antibody; scFv; human; H1b-1;
KW glycoprotein Ib alpha; platelet; aggregation; antiaggregant;
KW antithrombotic; thrombus; therapy; diagnostic.
OS Homo sapiens.

XX Key Location/Qualifiers
FH Region 1..22
FT Region /note= "framework region 1"
FT Region 23..35
FT Region /note= "complementarity determining region 1"
FT Region 34..48
FT Region /note= "framework region 2"
FT Region 49..55
FT Region /note= "complementarity determining region 2"
FT Region 56..87
FT Region /note= "framework region 3"
FT Region 88..96
FT Region /note= "complementarity determining region 3"
FT Region 97..107
FT Region /note= "framework region 4"
XX WO200026667-A1.
XX 11-MAY-2000.
XX 29-OCT-1999; 99WO-US025495.
XX 30-OCT-1998; 98US-0106275P.
XX (MILLER) MILLER J L.
XX Miller JL;
XX WPI; 2000-365744/31.
XX Isolated nucleic acid molecule encoding anti-human platelet glycoprotein
XX Ib alpha molecule useful for producing antibodies which inhibit platelet
XX aggregation.

XX Claim 18; Fig 5; 89pp; English.

XX The present sequence is that of the light chain variable region (VL) of
XX human single chain antibody (scFv) H1b-1 (see AAY95198), which is
XX directed against platelet glycoprotein Ib (GPIb). The H1b series of scFv
XX was isolated from a human synthetic VH and VL scFv library by 3 rounds of
XX phagemid selection against transfected CHO cells expressing the GPIb
XX alpha component of the GPIb/IX/V complex on their surface, followed by a
XX 4th round of selection against washed human platelets, and 2 final rounds
XX in which attempts were made to displace scFv from washed platelets by
XX flooding with murine monoclonal antibody or mimotope peptide (see
XX AAY95229). Whether displayed as surface proteins on a phagemid or
XX secreted as free scFv by *Escherichia coli*, the H1b scFv clones are
XX capable of inhibiting von Willebrand factor-dependent aggregation of
XX platelets. The scFv are composed of native human protein sequences and
XX are therefore attractive potential reagents for therapeutic purposes.
XX They provide a new class of antithrombotic agents, useful for the
XX prevention of platelet-dependent thrombi in diseased arteries, bypass
XX grafts, dialysis etc., and can also be used as diagnostic reagents.
XX Methods of inhibiting aggregation of platelets, of binding human platelet

CC GPIb alpha and of selecting a VH or VL region of an antibody that
XX inhibits platelet aggregation are claimed

XX Sequence 107 AA;
Query Match 87.7%; Score 50; DB 3; Length 107;
Best Local Similarity 81.8%; Pred. No. 0.21;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLSKTYVAS 11
Db 23 QGDSLSRYSYAS 33
|||:||||
|||:||||

RESULT 24
AAG62969

ID AAG62969 standard; protein; 107 AA.
XX AC AAG62969;
XX 01-OCT-2001 (first entry)
XX Amino acid sequence of variable light chain fragment of clone D5.
DE Antibody; light chain; VL; amyloid protein; blood brain barrier;
XX endothelial cell; brain cell antigen; inflammation; adhesion molecule;
KW transferrin receptor; neurological disease; Alzheimer's disease;
KW prion disease; AIDS-related dementia; epilepsy; brain injury.
XX Homo sapiens.
XX WO200144300-A2.
XX 21-JUN-2001.
XX 27-NOV-2000; 2000WO-GB004501.
XX 13-DEC-1999; 99US-0170599P.
XX (CAMP-) CAMBRIDGE ANTIBODY TECHNOLOGY.
XX Webster C, Osbourn J, Ward G, Miller K;
XX WPI; 2001-398131/42.
XX N-PSDB; AAH42412.
XX Mixture or panel of antibodies for selecting specific binding members
XX that cross the blood brain barrier, for use in delivering different
XX molecules and treating neurological diseases.
XX Claim 1; Page 109; 109pp; English.
XX The present sequence represents an antibody variable light chain (VL)
XX fragment. The fragment is used to produce a mixture or panel of 5
XX different specific binding members, each comprising an antibody VH and/or
XX VL variable domain and capable, when displayed on the surface of
XX filamentous bacteriophage particles or in the case of a specific binding
XX member comprising the D5 VH and/or VL variable domain when bound to human
XX serum amyloid protein, to pass through a mammalian blood brain barrier
XX (BBB). The panel is useful for the selection of specific binding members
XX with a desired property such as ability to cross BBB, ability to bind
XX endothelial cells or other brain cell antigen, ability to bind areas of
XX inflammation in the brain or BBB breakdown or ability to bind
XX intracellular adhesion molecules and to bind transferrin receptor. The
XX antibodies are useful in diagnosis, prophylaxis and treatment of human or
XX animal body, including neurological diseases, such as Alzheimer's
XX disease, prion disease, AIDS-related dementia, epilepsy and traumatic
XX brain injury and any diseases involving inflammation occurring within the
XX brain or central nervous system
XX Sequence 107 AA;
Query Match 87.7%; Score 50; DB 4; Length 107;

Best Local Similarity 81.8%; Pred. No. 0.21; Mismatches 2; Indels 0; Gaps 0;

Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
 Db 21 QGDSLRSYVAS 31
 |||||:|||||

RESULT 25
 AAY95185
 ID AAY95185 standard; protein; 108 AA.
 AC AAY95185;
 DT 29-AUG-2000 (first entry)
 XX
 DE Anti-platelet glycoprotein Ib human H1b-3 VL.
 KW Variable light chain; single chain antibody; scFv; human; H1b-3;
 KW glycoprotein Ib alpha; platelet; aggregation; antiaggregant;
 KW antithrombotic; thrombus; therapy; diagnostic.
 XX
 OS Homo sapiens.
 XX
 PH Key
 FT Region Location/Qualifiers
 FT 1..22 /note= "framework region 1"
 FT 23..33 /note= "complementarity determining region 1"
 FT 34..48 /note= "framework region 2"
 FT 49..55 /note= "complementarity determining region 2"
 FT 56..87 /note= "framework region 3"
 FT 88..96 /note= "complementarity determining region 3"
 FT 97..108 /note= "framework region 4"
 FT
 PN WO200026667-A1.
 XX
 PD 11-MAY-2000.
 XX
 PF 29-OCT-1999; 99WO-US025495.
 XX
 PR 30-OCT-1998; 98US-0106275P.
 XX
 PA (MILL/) MILLER J L.
 XX
 PI Miller JL;
 XX
 DR WPI; 2000-365744/31.
 DR N-PSDB; AAA27663.
 XX
 PT Isolated nucleic acid molecule encoding anti-human platelet glycoprotein
 PT Ib alpha molecule useful for producing antibodies which inhibit platelet
 PT aggregation.
 XX
 PS Claim 11; Page 76; 89pp; English.
 CC
 CC The present sequence is that of the light chain variable region (VL) of
 CC human single chain antibody (scFv) H1b-3 (see AAY95219), which is
 CC directed against platelet glycoprotein Ib (GPIb). The H1b series of scFv
 CC was isolated from a human synthetic VH and VL scFv library by 3 rounds of
 CC phagemid selection against transfected CHO cells expressing the GPIb
 CC alpha component of the GPIb/IX/V complex on their surface, followed by a
 CC 4th round of selection against washed human platelets, and 2 final rounds
 CC in which attempts were made to displace scFv from washed platelets by
 CC flooding with murine monoclonal antibody or mimotope peptide (see
 CC AAY95229). Whether displayed as surface proteins on a phagemid or
 CC secreted as free scFv by Escherichia coli, the H1b scFv clones are
 CC capable of inhibiting von Willebrand factor-dependent aggregation of

CC platelets. The scFv are composed of native human protein sequences and
 CC are therefore attractive potential reagents for therapeutic purposes.
 CC They provide a new class of antithrombotic agents, useful for the
 CC prevention of platelet-dependent thrombi in diseased arteries, bypass
 CC grafts, dialysis etc., and can also be used as diagnostic reagents.
 CC Methods of inhibiting aggregation of platelets, of binding human platelet
 CC GPIb alpha and of selecting a VH or VL region of an antibody that
 CC inhibits platelet aggregation are claimed

XX
 SQ Sequence 108 AA;
 Query Match 87.7%; Score 50; DB 3; Length 108;
 Best Local Similarity 81.8%; Pred. No. 0.21;
 Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
 Db 23 QGDSLRSYVAS 33
 |||||:|||||

Search completed: September 24, 2004, 01:52:58
 Job time : 74.7736 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 01:53:42 ; Search time 18.6792 Seconds
(without alignments)
30.402 Million cell updates/sec

Title: US-10-088-639A-2_COPY_23_33
Perfect score: 57
Sequence: 1 QGDSLKTYAS 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/prodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/prodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/prodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/prodata/2/iaa/PCUTUS_COMB.pep:*
6: /cgn2_6/prodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	54	94.7	104	4	US-08-793-450-2
2	54	94.7	238	4	US-08-793-450-6
3	50	87.7	97	2	US-08-665-202-35
4	50	87.7	97	4	US-09-315-574-35
5	50	87.7	103	2	US-08-273-146-71
6	50	87.7	104	3	US-09-240-274-49
7	50	87.7	109	2	US-08-652-816A-16
8	50	87.7	109	2	US-08-665-202-34
9	50	87.7	109	4	US-09-315-574-34
10	50	87.7	278	3	US-09-260-527-3
11	50	87.7	309	3	US-09-260-527-1
12	50	87.7	309	4	US-09-079-029-9
13	50	87.7	312	4	US-09-079-029-10
14	48	84.2	106	3	US-09-240-274-50
15	46	80.7	106	3	US-09-240-274-47
16	46	80.7	106	3	US-09-240-274-48
17	38	66.7	117	4	US-09-530-139-20
18	38	66.7	224	4	US-09-530-139-38
19	38	66.7	238	4	US-09-530-139-36
20	36	63.2	185	4	US-09-562-737-117
21	35	61.4	86	4	US-09-134-000C-5678
22	35	61.4	200	4	US-09-252-991A-19243
23	34	59.6	119	2	US-08-652-816A-10
24	34	59.6	244	4	US-08-518-148-79
25	34	59.6	376	4	US-09-328-352-4299
26	34	59.6	439	1	US-07-637-870-9
27	34	59.6	439	1	US-07-637-399-6

28	34	59.6	439	1	US-08-112-703-6	Sequence 6, Appli
29	34	59.6	657	3	US-08-893-852A-3	Sequence 3, Appli
30	34	59.6	657	3	US-08-821-818-3	Sequence 3, Appli
31	34	59.6	657	4	US-09-052-753B-3	Sequence 3, Appli
32	34	59.6	670	4	US-09-328-352-6725	Sequence 6725, Ap
33	33	57.9	100	4	US-09-899-896-12	Sequence 12, Appl
34	33	57.9	150	4	US-09-134-001C-5625	Sequence 5625, Ap
35	33	57.9	581	4	US-09-649-747A-13	Sequence 13, Appl
36	33	57.9	581	4	US-09-649-747A-21	Sequence 21, Appl
37	33	57.9	670	4	US-09-831-642-32	Sequence 32, Appl
38	32	56.1	108	4	US-09-025-769B-20	Sequence 20, Appl
39	32	56.1	207	4	US-09-328-352-6658	Sequence 6658, Ap
40	32	56.1	255	4	US-09-489-039A-13931	Sequence 13931, A
41	32	56.1	264	4	US-09-543-681A-4818	Sequence 4818, Ap
42	32	56.1	335	2	US-09-014-969-15	Sequence 15, Appl
43	32	56.1	405	4	US-09-071-035-4	Sequence 4, Appli
44	32	56.1	426	4	US-09-071-035-2	Sequence 2, Appli
45	32	56.1	436	4	US-09-393-858-11	Sequence 11, Appl
46	32	56.1	447	4	US-09-134-000C-5071	Sequence 5071, Ap
47	32	56.1	590	4	US-09-107-532A-4682	Sequence 4682, Ap
48	32	56.1	724	4	US-09-562-737-30	Sequence 30, Appl
49	32	56.1	725	4	US-09-134-000C-4702	Sequence 4702, Ap
50	32	56.1	765	1	US-08-309-512-5	Sequence 5, Appli
51	32	56.1	765	5	PCT-US92-08756A-5	Sequence 5, Appli
52	32	56.1	920	3	US-08-930-996A-8	Sequence 8, Appli
53	31.5	55.3	1439	4	US-09-543-681A-7560	Sequence 24, Appl
54	31	54.4	20	4	US-09-881-710-24	Sequence 8, Appli
55	31	54.4	50	5	PCT-US91-02942-8	Sequence 8, Appli
56	31	54.4	90	4	US-09-134-000C-5906	Sequence 5906, Ap
57	31	54.4	113	3	US-08-974-899-6	Sequence 6, Appli
58	31	54.4	119	4	US-09-648-067A-15	Sequence 15, Appl
59	31	54.4	122	2	US-07-934-373C-21	Sequence 21, Appl
60	31	54.4	122	3	US-08-437-642B-21	Sequence 21, Appl
61	31	54.4	122	4	US-08-146-206C-21	Sequence 21, Appl
62	31	54.4	122	4	US-09-705-686-21	Sequence 21, Appl
63	31	54.4	123	5	PCT-US93-07832-21	Sequence 21, Appl
64	31	54.4	123	1	US-08-199-911-4	Sequence 4, Appli
65	31	54.4	138	2	US-07-857-224B-112	Sequence 112, App
66	31	54.4	151	1	US-08-614-935-28	Sequence 28, Appl
67	31	54.4	151	3	US-09-130-287-28	Sequence 28, Appl
68	31	54.4	159	4	US-09-134-000C-3788	Sequence 3788, Ap
69	31	54.4	182	4	US-08-858-207A-434	Sequence 434, App
70	31	54.4	256	4	US-09-107-532A-3657	Sequence 3657, Ap
71	31	54.4	302	4	US-09-543-681A-4695	Sequence 4695, Ap
72	31	54.4	305	4	US-09-543-681A-4689	Sequence 4689, Ap
73	31	54.4	323	4	US-09-543-681A-4866	Sequence 4866, Ap
74	31	54.4	325	4	US-09-328-352-5100	Sequence 5100, Ap
75	31	54.4	328	4	US-09-134-001C-5167	Sequence 5167, Ap
76	31	54.4	337	4	US-09-252-991A-30261	Sequence 30261, A
77	31	54.4	345	4	US-09-543-681A-5409	Sequence 5409, Ap
78	31	54.4	355	2	US-08-834-655-5	Sequence 5, Appli
79	31	54.4	355	3	US-08-834-033A-6	Sequence 6, Appli
80	31	54.4	355	3	US-09-363-574-5	Sequence 5, Appli
81	31	54.4	355	4	US-09-363-526-5	Sequence 2, Appli
82	31	54.4	360	3	US-09-231-227-2	Sequence 2, Appli
83	31	54.4	360	4	US-09-768-585-2	Sequence 2, Appli
84	31	54.4	370	4	US-09-543-681A-6187	Sequence 6187, Ap
85	31	54.4	371	4	US-09-543-681A-7332	Sequence 7332, Ap
86	31	54.4	442	4	US-09-134-001C-3033	Sequence 3033, Ap
87	31	54.4	457	2	US-08-834-655-2	Sequence 2, Appli
88	31	54.4	457	2	US-08-833-610-4	Sequence 4, Appli
89	31	54.4	457	3	US-08-834-033A-2	Sequence 2, Appli
90	31	54.4	457	3	US-08-834-033A-14	Sequence 14, Appl
91	31	54.4	457	3	US-09-363-574-2	Sequence 2, Appli
92	31	54.4	457	4	US-09-363-526-2	Sequence 2, Appli
93	31	54.4	457	4	US-09-330-235-18	Sequence 18, Appl
94	31	54.4	457	4	US-09-377-452-4	Sequence 4, Appli
95	31	54.4	458	4	US-09-439-261-10	Sequence 10, Appl
96	31	54.4	458	4	US-09-439-261-44	Sequence 44, Appl
97	31	54.4	458	4	US-09-227-613-11	Sequence 11, Appl
98	31	54.4	458	4	US-09-227-613-41	Sequence 41, Appl
99	31	54.4	487	4	US-09-198-452A-907	Sequence 907, App
100	31	54.4	487	4	US-09-134-000C-6001	Sequence 6001, Ap

101 31 54.4 489 4 US-09-252-991A-19428 Sequence 19428, A
102 31 54.4 514 3 US-09-717-432-2 Sequence 2, Appli
103 31 54.4 514 3 US-09-912-484-2 Sequence 2, Appli
104 31 54.4 518 3 US-08-999-723-2 Sequence 2, Appli
105 31 54.4 518 3 US-09-434-427-2 Sequence 2, Appli
106 31 54.4 518 4 US-09-548-372D-2 Sequence 2, Appli
107 31 54.4 518 4 US-09-548-367D-2 Sequence 2, Appli
108 31 54.4 518 4 US-09-551-853D-2 Sequence 2, Appli
109 31 54.4 518 4 US-09-215-450-19 Sequence 19, Appl
110 31 54.4 614 5 PCT-US95-03236-21 Sequence 21, Appl
111 31 54.4 659 1 US-08-258-639A-4 Sequence 4, Appli
112 31 54.4 659 2 US-08-900-951-4 Sequence 4, Appli
113 31 54.4 659 5 PCT-US95-07391A-4 Sequence 4, Appli
114 31 54.4 795 4 US-09-198-452A-314 Sequence 314, App
115 31 54.4 893 4 US-09-514-308-4 Sequence 4, Appli
116 31 54.4 1101 4 US-09-561-709B-5 Sequence 5, Appli
117 31 54.4 1761 4 US-09-561-709B-1 Sequence 1, Appli
118 31 54.4 1938 4 US-09-514-302-2 Sequence 2, Appli
119 31 54.4 2285 4 US-09-308-375-2 Sequence 2, Appli
120 30.5 281 4 US-09-489-039A-9637 Sequence 9637, Ap
121 30.5 415 4 US-09-025-769B-280 Sequence 280, App
122 30.5 432 4 US-09-403-089A-1 Sequence 1, Appli
123 30.5 434 4 US-09-107-532A-5626 Sequence 5626, Ap
124 30.5 482 1 US-08-358-160-5 Sequence 5, Appli
125 30.5 484 1 US-08-358-160-7 Sequence 7, Appli
126 30 52.6 35 1 US-08-487-890A-67 Sequence 67, Appl
127 30 52.6 35 2 US-08-478-433-67 Sequence 67, Appl
128 30 52.6 35 2 US-08-337-483-67 Sequence 67, Appl
129 30 52.6 35 2 US-08-478-373-67 Sequence 67, Appl
130 30 52.6 35 3 US-08-474-671-67 Sequence 67, Appl
131 30 52.6 35 3 US-08-483-577A-67 Sequence 67, Appl
132 30 52.6 35 3 US-08-897-438-67 Sequence 67, Appl
133 30 52.6 35 4 US-08-637-654-67 Sequence 67, Appl
134 30 52.6 35 4 US-08-649-518-67 Sequence 67, Appl
135 30 52.6 39 4 US-08-857-046A-27 Sequence 27, Appl
136 30 52.6 66 4 US-09-634-238-336 Sequence 336, App
137 30 52.6 79 4 US-09-543-681A-6948 Sequence 6948, Ap
138 30 52.6 124 3 US-08-983-607-47 Sequence 47, Appl
139 30 52.6 137 1 US-08-468-709B-4 Sequence 4, Appli
140 30 52.6 137 1 US-08-468-709B-11 Sequence 11, Appl
141 30 52.6 137 2 US-08-241-664B-4 Sequence 4, Appli
142 30 52.6 137 2 US-08-241-664B-11 Sequence 11, Appl
143 30 52.6 137 5 PCT-US93-03936-4 Sequence 4, Appli
144 30 52.6 147 4 US-09-634-238-333 Sequence 333, App
145 30 52.6 191 4 US-09-134-000C-3705 Sequence 3705, Ap
146 30 52.6 210 4 US-09-107-532A-5401 Sequence 5401, Ap
147 30 52.6 210 4 US-09-134-000C-6797 Sequence 6797, Ap
148 30 52.6 227 1 US-08-244-646-17 Sequence 17, Appl
149 30 52.6 252 4 US-09-252-991A-29092 Sequence 29092, A
150 30 52.6 269 3 US-09-111-556A-4 Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-08-793-450-2
; Sequence 2, Application US/08793450
; Patent No. 6312690
; GENERAL INFORMATION:
; APPLICANT: EDELMAN, LENA
; APPLICANT: MARGARITTE, CHRISTEL
; APPLICANT: KACZOREK, MICHEL
; APPLICANT: CHAABIHI, HASSAN
; TITLE OF INVENTION: MONOCLONAL RECOMBINANT ANTI-RHESUS D
; TITLE OF INVENTION:
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLOH, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA

COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/793,450
FILING DATE: 03-MAR-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 94/10566
FILING DATE: 02-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: OBLOH, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 660-118-0 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 104 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-793-450-2
Query Match 94.7%; Score 54; DB 4; Length 104;
Best Local Similarity 90.9%; Pred. No. 0.0047;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 QGDSLKTYVAS 11
Db 23 QGDSLRTYVAS 33
RESULT 2
US-08-793-450-6
; Sequence 6, Application US/08793450
; Patent No. 6312690
; GENERAL INFORMATION:
; APPLICANT: EDELMAN, LENA
; APPLICANT: MARGARITTE, CHRISTEL
; APPLICANT: KACZOREK, MICHEL
; APPLICANT: CHAABIHI, HASSAN
; TITLE OF INVENTION: MONOCLONAL RECOMBINANT ANTI-RHESUS D
; TITLE OF INVENTION:
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLOH, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/793,450
FILING DATE: 03-MAR-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 94/10566
FILING DATE: 02-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: OBLOH, NORMAN F.
REGISTRATION NUMBER: 24,618

REFERENCE/DOCKET NUMBER: 660-118-0 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 238 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-793-450-6

Query Match 94.7%; Score 54; DB 4; Length 238;
Best Local Similarity 90.9%; Pred. No. 0.011;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
Db 42 QGDSLRTYVAS 52

RESULT 3

US-08-665-202-35
Sequence 35, Application US/08665202
Patent No. 5977322
GENERAL INFORMATION:
APPLICANT: Marks, James D.
APPLICANT: Schier, Robert
TITLE OF INVENTION: No. 5977322el High Affinity Human Antibodies to
TITLE OF INVENTION: Tumor Antigens
NUMBER OF SEQUENCES: 141
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/665,202
FILING DATE: 13-JUN-1996
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/000,238
FILING DATE: 14-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/000,250
FILING DATE: 15-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Hunter, Tom
REGISTRATION NUMBER: 38,498
REFERENCE/DOCKET NUMBER: 02307E-061410
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 97 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-665-202-35

Query Match 87.7%; Score 50; DB 2; Length 97;
Best Local Similarity 81.8%; Pred. No. 0.025;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
Db 22 QGDSLRSYVAS 32

RESULT 4

US-09-315-574-35
Sequence 35, Application US/09315574
Patent No. 6512097
GENERAL INFORMATION:
APPLICANT: Marks, James D.
APPLICANT: Schier, Robert
TITLE OF INVENTION: No. 6512097el High Affinity Human Antibodies to
TITLE OF INVENTION: Tumor Antigens
NUMBER OF SEQUENCES: 141
CORRESPONDENCE ADDRESS:
ADDRESSEE: Majestic, Parsons, Siebert & Heue P.C.
STREET: Four Embarcadero Center, Suite 1100
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4106
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/315,574
FILING DATE: 20-MAY-99
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/000,238
FILING DATE: 14-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/000,250
FILING DATE: 15-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/665,202
FILING DATE: 13-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Hunter, Tom
REGISTRATION NUMBER: 38,498
REFERENCE/DOCKET NUMBER: 02307E-061411
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 97 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-315-574-35

Query Match 87.7%; Score 50; DB 4; Length 97;
Best Local Similarity 81.8%; Pred. No. 0.025;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
Db 22 QGDSLRSYVAS 32

RESULT 5

US-08-273-146-71
Sequence 71, Application US/08273146
Patent No. 5855885
GENERAL INFORMATION:
APPLICANT: Smith, Rodger
APPLICANT: McCafferty, John
APPLICANT: Chiswell, David

APPLICANT: Darsley, Michael J.
APPLICANT: Fitzgerald, Kevin
APPLICANT: Kenten, John H.
APPLICANT: Martin, Mark T.
APPLICANT: Titmas, Richard C.
APPLICANT: Williams, Richard O.
TITLE OF INVENTION: The Isolation and Production of
TITLE OF INVENTION: Catalytic Antibodies using Phage Technology
NUMBER OF SEQUENCES: 71
CORRESPONDENCE ADDRESS:
ADDRESSEE: IGEN, Inc.
STREET: 1530 East Jefferson St.
CITY: Rockville
STATE: MD
COUNTRY: USA
ZIP: 20852
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION NUMBER: US/08/273,146
FILING DATE: 14-JUL-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ryan, John W.
REGISTRATION NUMBER: 33,771
REFERENCE/DOCKET NUMBER: 09000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-984-8000
TELEFAX: 301-230-0158
INFORMATION FOR SEQ ID NO: 71:
SEQUENCE CHARACTERISTICS:
LENGTH: 103 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-273-146-71
Query Match 87.7%; Score 50; DB 2; Length 103;
Best Local Similarity 81.8%; Pred. No. 0.027;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
|||||:|||||
Db 17 QGDSLRSYVAS 27

RESULT 6
US-09-240-274-49
Sequence 49, Application US/09240274
Patent No. 6255455
GENERAL INFORMATION:
APPLICANT: Siegel, Donald L.
TITLE OF INVENTION: Rh(D)-BINDING PROTEINS AND MAGNETICALLY ACTIVATED CELL
FILE REFERENCE: 09596-42U2
CURRENT APPLICATION NUMBER: US/09/240,274
CURRENT FILING DATE: 1999-01-29
EARLIER APPLICATION NUMBER: 60/081,380
EARLIER FILING DATE: 1998-04-10
EARLIER APPLICATION NUMBER: 60/028,550
EARLIER FILING DATE: 1996-10-11
NUMBER OF SEQ ID NOS: 224
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 49
LENGTH: 104
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: anti-Rh(D) chain J04
US-09-240-274-49

Query Match 87.7%; Score 50; DB 3; Length 104;
Best Local Similarity 81.8%; Pred. No. 0.027;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 QGDSLKTYVAS 11
|||||:|||||
Db 21 QGDSLRSYVAS 31

RESULT 7
US-08-652-816A-16
Sequence 16, Application US/08652816A
Patent No. 5872215
GENERAL INFORMATION:
APPLICANT: Osbourn, JK
APPLICANT: McCafferty, JG
TITLE OF INVENTION: Specific binding members, materials and
TITLE OF INVENTION: methods.
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION NUMBER: US/08/652,816A
FILING DATE: 23-MAY-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9125579.4
FILING DATE: 02-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9125579.8
FILING DATE: 02-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9206318.9
FILING DATE: 24-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9206372.6
FILING DATE: 23-SEP-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9525004.9
FILING DATE: 07-DEC-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9610824.6
FILING DATE: 23-MAY-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB92/02240
FILING DATE: 02-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/244,597
FILING DATE: 01-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: David W. Clough
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 28111/33308
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 109 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-652-816A-16

Query Match 87.7%; Score 50; DB 2; Length 109;

```
Best Local Similarity 81.8%; Pred. No. 0.029; Mismatches 0; Indels 0; Gaps 0;
Matches 9; Conservative 2;

QY 1 QGDSLKTYVAS 11
Db 23 QGDSLRSYYAS 33

RESULT 8
US-08-665-202-34
; Sequence 34, Application US/08665202
; Patent No. 5977322
; GENERAL INFORMATION:
; APPLICANT: Marks, James D.
; APPLICANT: Schier, Robert
; TITLE OF INVENTION: No. 5977322el High Affinity Human Antibodies to
; Tumor Antigens
; NUMBER OF SEQUENCES: 141
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/665,202
; FILING DATE: 13-JUN-1996
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/000,238
; FILING DATE: 14-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/000,250
; FILING DATE: 15-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Hunter, Tom
; REGISTRATION NUMBER: 38,498
; REFERENCE/DOCKET NUMBER: 02307E-061410
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 109 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-665-202-34

Query Match 87.7%; Score 50; DB 2; Length 109;
Best Local Similarity 81.8%; Pred. No. 0.029;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
Db 22 QGDSLRSYYAS 32

RESULT 9
US-09-315-574-34
; Sequence 34, Application US/09315574
; Patent No. 6512097
; GENERAL INFORMATION:
; APPLICANT: Marks, James D.
; APPLICANT: Schier, Robert
; TITLE OF INVENTION: No. 6512097el High Affinity Human Antibodies to
```

```
; Tumor Antigens
; NUMBER OF SEQUENCES: 141
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Majestic, Parsons, Siebert & Hsue P.C.
; STREET: Four Embarcadero Center, Suite 1100
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4106
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/315,574
; FILING DATE: 20-MAY-99
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/000,238
; FILING DATE: 14-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/000,250
; FILING DATE: 15-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/665,202
; FILING DATE: 13-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hunter, Tom
; REGISTRATION NUMBER: 38,498
; REFERENCE/DOCKET NUMBER: 02307E-061411
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 109 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-315-574-34

Query Match 87.7%; Score 50; DB 4; Length 109;
Best Local Similarity 81.8%; Pred. No. 0.029;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
Db 22 QGDSLRSYYAS 32

RESULT 10
US-09-260-527-3
; Sequence 3, Application US/09260527A
; Patent No. 6228599
; GENERAL INFORMATION:
; APPLICANT: Knox, J.P.
; APPLICANT: Mikkelson, J.D.
; APPLICANT: Willats, W. G.
; TITLE OF INVENTION: ANTIBODY
; FILE REFERENCE: DY0U19.001AUS
; CURRENT APPLICATION NUMBER: US/09/260,527A
; CURRENT FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 278
; TYPE: PPT
; ORGANISM: UNKNOWN
; FEATURE:
; OTHER INFORMATION: Anti-homogalacturonan specific antibodies selected
; OTHER INFORMATION: from a naive phage display library known as the
```

; OTHER INFORMATION: Synthetic scFv Library (#1) from the Centre for
; OTHER INFORMATION: Protein Engineering, MRC Centre, Cambridge, UK
US-09-260-527-3

Query Match 87.7%; Score 50; DB 3; Length 278;
Best Local Similarity 81.8%; Pred. No. 0.077;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
Db 175 QGDSLRSYYAS 185
|||||:|||||

RESULT 11

US-09-260-527-1
; Sequence 1, Application US/09260527A
; Patent No. 6228599
; GENERAL INFORMATION:
; APPLICANT: Koox, J.P.
; APPLICANT: Mikkelsen, J.D.
; APPLICANT: Willats, W. G.
; TITLE OF INVENTION: ANTIBODY
; FILE REFERENCE: DY0U19.001AUS
; CURRENT APPLICATION NUMBER: US/09/260,527A
; CURRENT FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 280
; TYPE: PRT
; ORGANISM: UNKNOWN
; FEATURE:
; OTHER INFORMATION: Anti-homogalacturonan specific antibodies from a
; OTHER INFORMATION: phage display library known as the Synthetic scFv
; OTHER INFORMATION: Library (#1) from the Centre for Protein
; OTHER INFORMATION: Engineering, MRC Centre, Cambridge, UK.
US-09-260-527-1

Query Match 87.7%; Score 50; DB 3; Length 280;
Best Local Similarity 81.8%; Pred. No. 0.078;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
Db 177 QGDSLRSYYAS 187
|||||:|||||

RESULT 12

US-09-079-029-9
; Sequence 9, Application US/09079029
; Patent No. 6342369
; GENERAL INFORMATION:
; APPLICANT: Adams, Camilla W.
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Chuntharapai, Anan
; APPLICANT: Kim, Kyung J.
; TITLE OF INVENTION: Apo-2 Receptor
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/079,029
; FILING DATE:

; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Marschang, Diane L.
; REGISTRATION NUMBER: 35,600
; REFERENCE/DOCKET NUMBER: P1101R2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-5416
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 309 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-09-079-029-9

Query Match 87.7%; Score 50; DB 4; Length 309;
Best Local Similarity 81.8%; Pred. No. 0.086;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
Db 197 QGDSLRSYYAS 207
|||||:|||||

RESULT 13

US-09-079-029-10
; Sequence 10, Application US/09079029
; Patent No. 6342369
; GENERAL INFORMATION:
; APPLICANT: Adams, Camilla W.
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Chuntharapai, Anan
; APPLICANT: Kim, Kyung J.
; TITLE OF INVENTION: Apo-2 Receptor
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/079,029
; FILING DATE:

; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Marschang, Diane L.
; REGISTRATION NUMBER: 35,600
; REFERENCE/DOCKET NUMBER: P1101R2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-5416
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 312 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-09-079-029-10

Query Match 87.7%; Score 50; DB 4; Length 312;
Best Local Similarity 81.8%; Pred. No. 0.087;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
Db 200 QGDSLRSYYAS 210
|||||:|||||

```
RESULT 14
US-09-240-274-50
; Sequence 50, Application US/09240274
; Patent No. 6255455
; GENERAL INFORMATION:
; APPLICANT: Siegel, Donald L.
; TITLE OF INVENTION: Rh(D)-BINDING PROTEINS AND MAGNETICALLY ACTIVATED CELL
; FILE REFERENCE: 09596-42U2
; CURRENT APPLICATION NUMBER: US/09/240,274
; CURRENT FILING DATE: 1999-01-29
; EARLIER APPLICATION NUMBER: 60/081,380
; EARLIER FILING DATE: 1998-04-10
; EARLIER APPLICATION NUMBER: 60/028,550
; EARLIER FILING DATE: 1996-10-11
; NUMBER OF SEQ ID NOS: 224
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 50
; LENGTH: 106
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: anti-Rh(D) chain J05
US-09-240-274-50

Query Match      84.2%; Score 48; DB 3; Length 106;
Best Local Similarity 81.8%; Pred. No. 0.067;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 QGDSLKTYTAS 11
Db      21 QGDGLRSYYAS 31
|||||: |||||

RESULT 15
US-09-240-274-47
; Sequence 47, Application US/09240274
; Patent No. 6255455
; GENERAL INFORMATION:
; APPLICANT: Siegel, Donald L.
; TITLE OF INVENTION: Rh(D)-BINDING PROTEINS AND MAGNETICALLY ACTIVATED CELL
; FILE REFERENCE: 09596-42U2
; CURRENT APPLICATION NUMBER: US/09/240,274
; CURRENT FILING DATE: 1999-01-29
; EARLIER APPLICATION NUMBER: 60/081,380
; EARLIER FILING DATE: 1998-04-10
; EARLIER APPLICATION NUMBER: 60/028,550
; EARLIER FILING DATE: 1996-10-11
; NUMBER OF SEQ ID NOS: 224
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 47
; LENGTH: 106
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: anti-Rh(D) chain J01
US-09-240-274-47

Query Match      80.7%; Score 46; DB 3; Length 106;
Best Local Similarity 72.7%; Pred. No. 0.16;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 QGDSLKTYTAS 11
Db      21 QGDGLRSYYAS 31
|||||: |||||

RESULT 16
US-09-240-274-48
; Sequence 48, Application US/09240274
; Patent No. 6255455
; GENERAL INFORMATION:
; APPLICANT: Siegel, Donald L.
; TITLE OF INVENTION: Rh(D)-BINDING PROTEINS AND MAGNETICALLY ACTIVATED CELL
; FILE REFERENCE: 09596-42U2
; CURRENT APPLICATION NUMBER: US/09/240,274
; CURRENT FILING DATE: 1999-01-29
; EARLIER APPLICATION NUMBER: 60/081,380
; EARLIER FILING DATE: 1998-04-10
; EARLIER APPLICATION NUMBER: 60/028,550
; EARLIER FILING DATE: 1996-10-11
; NUMBER OF SEQ ID NOS: 224
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 48
; LENGTH: 106
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: anti-Rh(D) chain J02
US-09-240-274-48

Query Match      80.7%; Score 46; DB 3; Length 106;
Best Local Similarity 72.7%; Pred. No. 0.16;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 QGDSLKTYTAS 11
Db      21 QGDGLRSYYAS 31
|||||: |||||

RESULT 17
US-09-530-139-20
; Sequence 20, Application US/09530139
; Patent No. 6670453
; GENERAL INFORMATION:
; APPLICANT: FRENKEN, LEON GERARDUS
; APPLICANT: HOWELL, STEVEN
; APPLICANT: LEDEBOER, ADRIANUS MARINUS
; APPLICANT: VAN DER LOGT, CORNELIS PAUL
; TITLE OF INVENTION: MULTIVALENT ANTIGEN-BINDING PROTEINS
; FILE REFERENCE: 60113/268075/ASH
; CURRENT APPLICATION NUMBER: US/09/530,139
; CURRENT FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: PCT/EP98/06991
; PRIOR FILING DATE: 1998-10-27
; PRIOR APPLICATION NUMBER: EP 97308538.4
; PRIOR FILING DATE: 1997-10-27
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Lama peruana
US-09-530-139-20

Query Match      66.7%; Score 38; DB 4; Length 117;
Best Local Similarity 77.8%; Pred. No. 6;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      3 DSLKTYTAS 11
Db      54 DSARTYYAS 62
|||||: |||||

RESULT 18
US-09-530-139-38
; Sequence 38, Application US/09530139
; Patent No. 6670453
; GENERAL INFORMATION:
; APPLICANT: FRENKEN, LEON GERARDUS
; APPLICANT: HOWELL, STEVEN
; APPLICANT: LEDEBOER, ADRIANUS MARINUS
; APPLICANT: VAN DER LOGT, CORNELIS PAUL
```

; TITLE OF INVENTION: MULTIVALENT ANTIGEN-BINDING PROTEINS
; FILE REFERENCE: 60113/268075/ASH
; CURRENT APPLICATION NUMBER: US/09/530,139
; CURRENT FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: PCT/EP98/06991
; PRIOR FILING DATE: 1998-10-27
; PRIOR APPLICATION NUMBER: EP 97308538.4
; PRIOR FILING DATE: 1997-10-27
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 38
; LENGTH: 224
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Bispecific biheaded
; OTHER INFORMATION: antigen binding protein
US-09-530-139-38

Query Match 66.7%; Score 38; DB 4; Length 224;
Best Local Similarity 77.8%; Pred. No. 12;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 DSLKTYIAS 11
||:|||||
Db 50 DSARTYYAS 58

RESULT 19
US-09-530-139-36
; Sequence 36, Application US/09530139
; Patent No. 6670453
; GENERAL INFORMATION:
; APPLICANT: FRENKEN, LEON GERARDUS
; APPLICANT: HOWELL, STEVEN
; APPLICANT: LEDEBOER, ADRIANUS MARINUS
; APPLICANT: VAN DER LOGT, CORNELIS PAUL
; TITLE OF INVENTION: MULTIVALENT ANTIGEN-BINDING PROTEINS
; FILE REFERENCE: 60113/268075/ASH
; CURRENT APPLICATION NUMBER: US/09/530,139
; CURRENT FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: PCT/EP98/06991
; PRIOR FILING DATE: 1998-10-27
; PRIOR APPLICATION NUMBER: EP 97308538.4
; PRIOR FILING DATE: 1997-10-27
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 36
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Bispecific biheaded
; OTHER INFORMATION: antigen binding protein
US-09-530-139-36

Query Match 66.7%; Score 38; DB 4; Length 238;
Best Local Similarity 77.8%; Pred. No. 13;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 DSLKTYIAS 11
||:|||||
Db 50 DSARTYYAS 58

RESULT 20
US-09-562-737-117
; Sequence 117, Application US/09562737
; Patent No. 6428967
; GENERAL INFORMATION:
; APPLICANT: Herz, Joachim
; APPLICANT: Gotthardt, Michael
; TITLE OF INVENTION: LDL Receptor Signaling Pathways

; FILE REFERENCE: UTSW0708
; CURRENT APPLICATION NUMBER: US/09/562,737
; CURRENT FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 117
; LENGTH: 185
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Sequence
US-09-562-737-117

Query Match 63.2%; Score 36; DB 4; Length 185;
Best Local Similarity 54.5%; Pred. No. 23;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 QGDSLKTYIAS 11
:||||:|
Db 46 RGNLETYWQS 56

RESULT 21
US-09-134-000C-5678
; Sequence 5678, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 5678
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-5678

Query Match 61.4%; Score 35; DB 4; Length 86;
Best Local Similarity 66.7%; Pred. No. 16;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 QGDSLKTYIY 9
:||||:|
Db 1 EGD SMKTYIY 9

RESULT 22
US-09-252-991A-19243
; Sequence 19243, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19243
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa

Query Match 61.4%; Score 35; DB 4; Length 86;
Best Local Similarity 66.7%; Pred. No. 16;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 QGDSLKTYIY 9
:||||:|
Db 1 EGD SMKTYIY 9

US-09-252-991A-19243

Query Match 61.4%; Score 35; DB 4; Length 200;
Best Local Similarity 66.7%; Pred. No. 39;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 QGSLKTYT 9
|||: |||
Db 104 QGDLPLRDY 112

RESULT 23

US-08-652-816A-10
; Sequence 10, Application US/08652816A
; Patent No. 5872215
; GENERAL INFORMATION:
; APPLICANT: Osbourn, JK
; APPLICANT: Allen, DJ
; APPLICANT: McCafferty, JG
; TITLE OF INVENTION: Specific binding members, materials and
; TITLE OF INVENTION: methods.
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/652,816A
; FILING DATE: 23-MAY-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9125579.4
; FILING DATE: 02-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9125579.8
; FILING DATE: 02-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9206318.9
; FILING DATE: 24-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9206372.6
; FILING DATE: 23-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9525004.9
; FILING DATE: 07-DEC-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9610824.6
; FILING DATE: 23-MAY-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB92/02240
; FILING DATE: 02-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/244,597
; FILING DATE: 01-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: David W. Clough
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 28111/33308
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 119 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

US-08-652-816A-10

Query Match 59.6%; Score 34; DB 2; Length 119;
Best Local Similarity 60.0%; Pred. No. 35;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 GDSLKTYTAS 11
|||: |||
Db 26 GDSISSYYWS 35

RESULT 24

US-08-918-148-79
; Sequence 79, Application US/08918148A
; Patent No. 6342220
; GENERAL INFORMATION:
; APPLICANT: Adams, Camellia
; APPLICANT: W.
; APPLICANT: Carter, Paul J.
; APPLICANT: Fendly, Brian M.
; APPLICANT: Gurney, Austin L.
; TITLE OF INVENTION: Agonist Antibodies
; FILE REFERENCE: P0979
; CURRENT APPLICATION NUMBER: US/08/918,148A
; CURRENT FILING DATE: 1997-08-25
; NUMBER OF SEQ ID NOS: 79
; SEQ ID NO 79
; LENGTH: 244
; TYPE: PRT
; ORGANISM: artificial
US-08-918-148-79

Query Match 59.6%; Score 34; DB 4; Length 244;
Best Local Similarity 60.0%; Pred. No. 75;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 GDSLKTYTAS 11
|||: |||
Db 28 GDSISSYYWS 37

RESULT 25

US-09-328-352-4299
; Sequence 4299, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4299
; LENGTH: 376
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-4299

Query Match 59.6%; Score 34; DB 4; Length 376;
Best Local Similarity 77.8%; Pred. No. 1.2e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 DSLKTYTAS 11
|||: |||
Db 289 DALKTYAAS 297

Search completed: September 24, 2004, 02:19:55
Job time : 21.6792 secs

This Page Blank (uspto)


```

89 50 87.7 242 10 US-09-880-748-1424
90 50 87.7 242 10 US-09-880-748-1433
91 50 87.7 242 10 US-09-880-748-1444
92 50 87.7 242 10 US-09-880-748-1478
93 50 87.7 242 10 US-09-880-748-1762
94 50 87.7 242 10 US-09-880-748-2021
95 50 87.7 242 10 US-09-880-748-2046
96 50 87.7 242 10 US-09-880-748-2106
97 50 87.7 242 12 US-10-293-418-1424
98 50 87.7 242 12 US-10-293-418-1433
99 50 87.7 242 12 US-10-293-418-1444
100 50 87.7 242 12 US-10-293-418-1678
101 50 87.7 242 12 US-10-293-418-1762
102 50 87.7 242 12 US-10-293-418-2021
103 50 87.7 242 12 US-10-293-418-2046
104 50 87.7 242 12 US-10-293-418-2106
105 50 87.7 243 10 US-09-880-748-1929
106 50 87.7 243 10 US-09-880-748-1940
107 50 87.7 243 10 US-09-880-748-1942
108 50 87.7 243 10 US-09-880-748-1944
109 50 87.7 243 10 US-09-880-748-1947
110 50 87.7 243 10 US-09-880-748-2014
111 50 87.7 243 10 US-09-880-748-2056
112 50 87.7 243 12 US-09-880-748-2107
113 50 87.7 243 12 US-10-293-418-1929
114 50 87.7 243 12 US-10-293-418-1940
115 50 87.7 243 12 US-10-293-418-1942
116 50 87.7 243 12 US-10-293-418-1944
117 50 87.7 243 12 US-10-293-418-1947
118 50 87.7 243 12 US-10-293-418-2014
119 50 87.7 243 12 US-10-293-418-2056
120 50 87.7 243 12 US-10-293-418-2107
121 50 87.7 244 10 US-09-880-748-1692
122 50 87.7 244 10 US-09-880-748-1694
123 50 87.7 244 10 US-09-880-748-1694
124 50 87.7 244 10 US-09-880-748-2058
125 50 87.7 244 12 US-10-293-418-1692
126 50 87.7 244 12 US-10-293-418-1694
127 50 87.7 244 12 US-10-293-418-1924
128 50 87.7 244 12 US-10-293-418-2058
129 50 87.7 244 14 US-10-322-673-42
130 50 87.7 244 14 US-10-322-673-45
131 50 87.7 245 10 US-09-880-748-1615
132 50 87.7 245 10 US-09-880-748-1826
133 50 87.7 245 10 US-09-880-748-1919
134 50 87.7 245 12 US-10-293-418-1615
135 50 87.7 245 12 US-10-293-418-1826
136 50 87.7 245 12 US-10-293-418-1919
137 50 87.7 245 14 US-10-322-673-43
138 50 87.7 246 10 US-09-880-748-928
139 50 87.7 246 10 US-09-880-748-1192
140 50 87.7 246 10 US-09-880-748-1300
141 50 87.7 246 10 US-09-880-748-1308
142 50 87.7 246 10 US-09-880-748-1426
143 50 87.7 246 10 US-09-880-748-1655
144 50 87.7 246 10 US-09-880-748-1702
145 50 87.7 246 10 US-09-880-748-1818
146 50 87.7 246 10 US-09-880-748-1975
147 50 87.7 246 12 US-10-293-418-928
148 50 87.7 246 12 US-10-293-418-1192
149 50 87.7 246 12 US-10-293-418-1300
150 50 87.7 246 12 US-10-293-418-1308

```

ALIGNMENTS

```

RESULT 1
US-09-880-748-1530
; Sequence 1530, Application US/09880748
; Publication No. US20030059937A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.

```

```

; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PFS23
; CURRENT APPLICATION NUMBER: US/09/880,748
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 3239
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1530
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-748-1530

```

```

Query Match          93.0%; Score 53; DB 10; Length 253;
Best Local Similarity 90.9%; Pred. No. 0.096;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 QGDSLKTYVAS 11
        |||||:||||
Db      166 QGDSLKSYVAS 176

```

RESULT 2

```

US-10-293-418-1530
; Sequence 1530, Application US/10293418
; Publication No. US20030223996A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PFS23P2
; CURRENT APPLICATION NUMBER: US/10/293,418
; CURRENT FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: 60/331,469
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/340,817
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 09/880,748
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1530
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-418-1530

```

```

Query Match          93.0%; Score 53; DB 12; Length 253;
Best Local Similarity 90.9%; Pred. No. 0.096;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 QGDSLKTYVAS 11
        |||||:||||
Db      166 QGDSLKSYVAS 176

```

RESULT 3

US-09-880-748-2084
; Sequence 2084, Application US/09880748
; Publication No. US20030059937A1
; GENERAL INFORMATION:

; APPLICANT: Ruben et al.

; TITLE OF INVENTION: Antibodies that Immunospécifically Bind Blys

; FILE REFERENCE: PF523

; CURRENT APPLICATION NUMBER: US/09/880,748

; CURRENT FILING DATE: 2001-06-15

; PRIOR APPLICATION NUMBER: 60/212,210

; PRIOR FILING DATE: 2000-06-15

; PRIOR APPLICATION NUMBER: 60/240,816

; PRIOR FILING DATE: 2000-10-17

; PRIOR APPLICATION NUMBER: 60/276,248

; PRIOR FILING DATE: 2001-03-16

; PRIOR APPLICATION NUMBER: 60/277,379

; PRIOR FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/293,499

; PRIOR FILING DATE: 2001-05-25

; NUMBER OF SEQ ID NOS: 3239

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 2084

; LENGTH: 242

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-880-748-2084

Query Match 89.5%; Score 51; DB 10; Length 242;

Best Local Similarity 81.8%; Pred. No. 0.22;

Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11

DB 156 QGDSLRTYYAN 166

|||||:||||:

RESULT 4

US-10-293-418-2084
; Sequence 2084, Application US/10293418
; Publication No. US20030223996A1
; GENERAL INFORMATION:

; APPLICANT: Ruben et al.

; TITLE OF INVENTION: Antibodies that Immunospécifically Bind Blys

; FILE REFERENCE: PF523P2

; CURRENT APPLICATION NUMBER: US/10/293,418

; CURRENT FILING DATE: 2002-11-27

; PRIOR APPLICATION NUMBER: 60/331,469

; PRIOR FILING DATE: 2001-11-16

; PRIOR APPLICATION NUMBER: 60/340,817

; PRIOR FILING DATE: 2001-12-19

; PRIOR APPLICATION NUMBER: 09/880,748

; PRIOR FILING DATE: 2001-06-15

; PRIOR APPLICATION NUMBER: 60/293,499

; PRIOR FILING DATE: 2001-05-25

; PRIOR APPLICATION NUMBER: 60/277,379

; PRIOR FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/276,248

; PRIOR FILING DATE: 2001-03-16

; PRIOR APPLICATION NUMBER: 60/240,816

; PRIOR FILING DATE: 2000-10-17

; PRIOR APPLICATION NUMBER: 60/212,210

; PRIOR FILING DATE: 2000-06-16

; NUMBER OF SEQ ID NOS: 3247

; SEQ ID NO 2084

; LENGTH: 242

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-293-418-2084

Query Match

Best Local Similarity 89.5%; Score 51; DB 12; Length 242;

Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11

DB 156 QGDSLRTYYAN 166

|||||:||||:

RESULT 5

US-09-880-748-2102

; Sequence 2102, Application US/09880748

; Publication No. US20030059937A1

; GENERAL INFORMATION:

; APPLICANT: Ruben et al.

; TITLE OF INVENTION: Antibodies that Immunospécifically Bind Blys

; FILE REFERENCE: PF523

; CURRENT APPLICATION NUMBER: US/09/880,748

; CURRENT FILING DATE: 2001-06-15

; PRIOR APPLICATION NUMBER: 60/212,210

; PRIOR FILING DATE: 2000-06-15

; PRIOR APPLICATION NUMBER: 60/240,816

; PRIOR FILING DATE: 2000-10-17

; PRIOR APPLICATION NUMBER: 60/276,248

; PRIOR FILING DATE: 2001-03-16

; PRIOR APPLICATION NUMBER: 60/277,379

; PRIOR FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/293,499

; PRIOR FILING DATE: 2001-05-25

; NUMBER OF SEQ ID NOS: 3239

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 2102

; LENGTH: 243

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-880-748-2102

Query Match

Best Local Similarity 89.5%; Score 51; DB 10; Length 243;

Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11

DB 157 QGDSLRTYYAN 167

|||||:||||:

RESULT 6

US-10-293-418-2102

; Sequence 2102, Application US/10293418

; Publication No. US20030223996A1

; GENERAL INFORMATION:

; APPLICANT: Ruben et al.

; TITLE OF INVENTION: Antibodies that Immunospécifically Bind Blys

; FILE REFERENCE: PF523P2

; CURRENT APPLICATION NUMBER: US/10/293,418

; CURRENT FILING DATE: 2002-11-27

; PRIOR APPLICATION NUMBER: 60/331,469

; PRIOR FILING DATE: 2001-11-16

; PRIOR APPLICATION NUMBER: 60/340,817

; PRIOR FILING DATE: 2001-12-19

; PRIOR APPLICATION NUMBER: 09/880,748

; PRIOR FILING DATE: 2001-06-15

; PRIOR APPLICATION NUMBER: 60/293,499

; PRIOR FILING DATE: 2001-05-25

; PRIOR APPLICATION NUMBER: 60/277,379

; PRIOR FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/276,248

; PRIOR FILING DATE: 2001-03-16

; PRIOR APPLICATION NUMBER: 60/240,816

; PRIOR FILING DATE: 2000-10-17

; PRIOR APPLICATION NUMBER: 60/212,210

; PRIOR FILING DATE: 2000-06-16

; NUMBER OF SEQ ID NOS: 3247

; SEQ ID NO 2102

; LENGTH: 243

; TYPE: PRT

```

; ORGANISM: Homo sapiens
US-10-293-418-2102

Query Match      89.5%; Score 51; DB 12; Length 243;
Best Local Similarity 81.8%; Pred. No. 0.22;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
   |||||:||||:
Db 157 QGDSLRTYYAN 167

RESULT 7
US-09-880-748-2103
; Sequence 2103, Application US/09880748
; Publication No. US20030059937A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523
; CURRENT APPLICATION NUMBER: US/09/880,748
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 3239
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2103
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-748-2103

Query Match      89.5%; Score 51; DB 10; Length 247;
Best Local Similarity 81.8%; Pred. No. 0.22;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
   |||||:||||:
Db 161 QGDSLRTYYAN 171

RESULT 8
US-10-293-418-2103
; Sequence 2103, Application US/10293418
; Publication No. US20030223996A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523P2
; CURRENT APPLICATION NUMBER: US/10/293,418
; CURRENT FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: 60/331,469
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/340,817
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 09/880,748
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2103
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-748-2103

Query Match      89.5%; Score 51; DB 10; Length 247;
Best Local Similarity 81.8%; Pred. No. 0.22;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
   |||||:||||:
Db 161 QGDSLRTYYAN 171

RESULT 9
US-09-880-748-2088
; Sequence 2088, Application US/09880748
; Publication No. US20030059937A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523
; CURRENT APPLICATION NUMBER: US/09/880,748
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 3239
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2088
; LENGTH: 248
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-748-2088

Query Match      89.5%; Score 51; DB 10; Length 248;
Best Local Similarity 81.8%; Pred. No. 0.22;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
   |||||:||||:
Db 162 QGDSLRTYYAN 172

RESULT 10
US-10-293-418-2088
; Sequence 2088, Application US/10293418
; Publication No. US20030223996A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523P2
; CURRENT APPLICATION NUMBER: US/10/293,418
; CURRENT FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: 60/331,469
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/340,817
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 09/880,748
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2088
; LENGTH: 248
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-748-2088

```

```
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 2088
; LENGTH: 248
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-418-2002

Query Match      89.5%; Score 51; DB 12; Length 248;
Best Local Similarity 81.8%; Pred. No. 0.22;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
Db 162 QGDSLRTYYAN 172

RESULT 11
US-09-880-748-2002
; Sequence 2002, Application US/09880748
; Publication No. US20030059937A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523
; CURRENT APPLICATION NUMBER: US/09/880,748
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 3239
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2002
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-748-2002

Query Match      89.5%; Score 51; DB 10; Length 249;
Best Local Similarity 81.8%; Pred. No. 0.22;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
Db 163 QGDSLRTYYAN 173

RESULT 12
US-10-293-418-2002
; Sequence 2002, Application US/10293418
; Publication No. US20030223996A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523P2
; CURRENT APPLICATION NUMBER: US/10/293,418
; CURRENT FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: 60/331,469
; PRIOR FILING DATE: 2001-11-16

; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 09/880,748
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 2002
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-418-2002

Query Match      89.5%; Score 51; DB 12; Length 249;
Best Local Similarity 81.8%; Pred. No. 0.22;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
Db 163 QGDSLRTYYAN 173

RESULT 13
US-09-880-748-1990
; Sequence 1990, Application US/09880748
; Publication No. US20030059937A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523
; CURRENT APPLICATION NUMBER: US/09/880,748
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 3239
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1990
; LENGTH: 250
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-748-1990

Query Match      89.5%; Score 51; DB 10; Length 250;
Best Local Similarity 81.8%; Pred. No. 0.22;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
Db 164 QGDSLRTYYAN 174

RESULT 14
US-10-293-418-1990
; Sequence 1990, Application US/10293418
; Publication No. US20030223996A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
```

; TITLE OF INVENTION: Antibodies that Immunospecifically Bind BlyS

; FILE REFERENCE: PFS23P2

; CURRENT APPLICATION NUMBER: US/10/293,418

; CURRENT FILING DATE: 2002-11-27

; PRIOR APPLICATION NUMBER: 60/331,469

; PRIOR FILING DATE: 2001-11-16

; PRIOR APPLICATION NUMBER: 60/340,817

; PRIOR FILING DATE: 2001-12-19

; PRIOR APPLICATION NUMBER: 09/880,748

; PRIOR FILING DATE: 2001-06-15

; PRIOR APPLICATION NUMBER: 60/293,499

; PRIOR FILING DATE: 2001-05-25

; PRIOR APPLICATION NUMBER: 60/277,379

; PRIOR FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/276,248

; PRIOR FILING DATE: 2001-03-16

; PRIOR APPLICATION NUMBER: 60/240,816

; PRIOR FILING DATE: 2000-10-17

; PRIOR APPLICATION NUMBER: 60/212,210

; PRIOR FILING DATE: 2000-06-16

; NUMBER OF SEQ ID NOS: 3247

; SEQ ID NO 1990

; LENGTH: 250

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-293-418-1990

Query Match 89.5%; Score 51; DB 12; Length 250;

Best Local Similarity 81.8%; Pred. No. 0.22;

Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11

DB 164 QGDSLRTYYAN 174

RESULT 15

US-10-453-698-106

; Sequence 106, Application US/10453698

; Publication No. US2004038308A1

; GENERAL INFORMATION:

; APPLICANT: Rother, Russell

; TITLE OF INVENTION: HYBRID ANTIBODIES

; FILE REFERENCE: 82 CIP (1087-37 CIP)

; CURRENT APPLICATION NUMBER: US/10/453,698

; NUMBER OF SEQ ID NOS: 196

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 106

; LENGTH: 96

; TYPE: PRT

; ORGANISM: human

US-10-453-698-106

Query Match 87.7%; Score 50; DB 12; Length 96;

Best Local Similarity 81.8%; Pred. No. 0.13;

Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11

DB 23 QGDSLRSYYAS 33

RESULT 16

US-10-308-817-106

; Sequence 106, Application US/10308817

; Publication No. US20030219861A1

; GENERAL INFORMATION:

; APPLICANT: Rother, Russell

; APPLICANT: Wu, Dayang

; TITLE OF INVENTION: HYBRID ANTIBODIES

; FILE REFERENCE: 1087-37

; CURRENT APPLICATION NUMBER: US/10/308,817

; CURRENT FILING DATE: 2002-12-03

; NUMBER OF SEQ ID NOS: 195

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 106

; LENGTH: 96

; TYPE: PRT

; ORGANISM: human

US-10-308-817-106

Query Match 87.7%; Score 50; DB 15; Length 96;

Best Local Similarity 81.8%; Pred. No. 0.13;

Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11

DB 23 QGDSLRSYYAS 33

RESULT 17

US-10-803-622-168

; Sequence 168, Application US/10803622

; Publication No. US20040157214A1

; GENERAL INFORMATION:

; APPLICANT: Cambridge Antibody Technology

; APPLICANT: Medical Research Council

; APPLICANT: McCafferty, John

; APPLICANT: Pope, Anthony

; APPLICANT: Johnson, Kevin

; APPLICANT: Hoogenboom, Hendricus

; APPLICANT: Griffiths, Andrew

; APPLICANT: Jackson, Ronald

; APPLICANT: Holliger, Kasper

; APPLICANT: Marks, James

; APPLICANT: Clackson, Timothy

; APPLICANT: Chiswell, David

; APPLICANT: Winter, Gregory

; APPLICANT: Bonert, Timothy

; TITLE OF INVENTION: Methods for Producing Members of Specific Binding Pairs

; FILE REFERENCE: 13839-00013

; CURRENT APPLICATION NUMBER: US/10/803,622

; CURRENT FILING DATE: 2004-03-18

; PRIOR APPLICATION NUMBER: GB 9015198.6

; PRIOR FILING DATE: 1990-07-10

; PRIOR APPLICATION NUMBER: GB 9022845.3

; PRIOR FILING DATE: 1990-10-19

; PRIOR APPLICATION NUMBER: GB 9022845.3

; PRIOR FILING DATE: 1990-10-19

; PRIOR APPLICATION NUMBER: GB 9024503.6

; PRIOR FILING DATE: 1990-11-12

; PRIOR APPLICATION NUMBER: GB 9104744.9

; PRIOR FILING DATE: 1991-03-06

; PRIOR APPLICATION NUMBER: GB 9110549.4

; PRIOR FILING DATE: 1991-05-15

; PRIOR APPLICATION NUMBER: PCT/GB91/01134

; PRIOR FILING DATE: 1991-07-10

; PRIOR APPLICATION NUMBER: US 07/971,857

; PRIOR FILING DATE: 1993-01-08

; PRIOR APPLICATION NUMBER: US 08/484,893

; PRIOR FILING DATE: 1995-06-07

; NUMBER OF SEQ ID NOS: 272

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 168

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-803-622-168

Query Match 87.7%; Score 50; DB 16; Length 101;

Best Local Similarity 81.8%; Pred. No. 0.13;

Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11

7.
8.
9.
10.
11.

Query Match 87.7%; Score 50; DB 14; Length 111;
Best Local Similarity 81.8%; Pred. No. 0.15;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
Db 23 QGDSLSRYVAS 33

RESULT 22

US-10-300-675-16
; Sequence 16, Application US/10300675
; Publication No. US20030198638A1

; GENERAL INFORMATION:
; APPLICANT: Watkins, Jeffrey D.
; TITLE OF INVENTION: Tumor Specific Monoclonal Antibodies
; FILE REFERENCE: P-IX 5519
; CURRENT APPLICATION NUMBER: US/10/300,675
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: US 09/989,901
; PRIOR FILING DATE: 2001-11-19
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Recombinant variant
US-10-300-675-16

Query Match 87.7%; Score 50; DB 14; Length 111;
Best Local Similarity 81.8%; Pred. No. 0.15;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
Db 23 QGDSLSRYVAS 33

RESULT 23

US-10-300-675-18
; Sequence 18, Application US/10300675
; Publication No. US20030198638A1

; GENERAL INFORMATION:
; APPLICANT: Watkins, Jeffrey D.
; TITLE OF INVENTION: Tumor Specific Monoclonal Antibodies
; FILE REFERENCE: P-IX 5519
; CURRENT APPLICATION NUMBER: US/10/300,675
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: US 09/989,901
; PRIOR FILING DATE: 2001-11-19
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Recombinant variant
US-10-300-675-18

Query Match 87.7%; Score 50; DB 14; Length 111;
Best Local Similarity 81.8%; Pred. No. 0.15;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
Db 23 QGDSLSRYVAS 33

RESULT 24

US-10-300-675-20

; Sequence 20, Application US/10300675
; Publication No. US20030198638A1

; GENERAL INFORMATION:
; APPLICANT: Watkins, Jeffrey D.
; TITLE OF INVENTION: Tumor Specific Monoclonal Antibodies
; FILE REFERENCE: P-IX 5519
; CURRENT APPLICATION NUMBER: US/10/300,675
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: US 09/989,901
; PRIOR FILING DATE: 2001-11-19
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Recombinant variant
US-10-300-675-20

Query Match 87.7%; Score 50; DB 14; Length 111;
Best Local Similarity 81.8%; Pred. No. 0.15;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
Db 23 QGDSLSRYVAS 33

RESULT 25

US-10-300-675-22
; Sequence 22, Application US/10300675
; Publication No. US20030198638A1

; GENERAL INFORMATION:
; APPLICANT: Watkins, Jeffrey D.
; TITLE OF INVENTION: Tumor Specific Monoclonal Antibodies
; FILE REFERENCE: P-IX 5519
; CURRENT APPLICATION NUMBER: US/10/300,675
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: US 09/989,901
; PRIOR FILING DATE: 2001-11-19
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Recombinant variant
US-10-300-675-22

Query Match 87.7%; Score 50; DB 14; Length 111;
Best Local Similarity 81.8%; Pred. No. 0.15;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
Db 23 QGDSLSRYVAS 33

Search completed: September 24, 2004, 02:25:07
Job time : 64.717 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 01:33:11 ; Search time 16.3962 Seconds
(without alignments)
64.534 Million cell updates/sec

Title: US-10-088-639a-2_COPY_23_33
Perfect score: 57
Sequence: 1 QGDSIKTYVAS 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Maximum Match 0%
Listing first 150 summaries

Database : PIR 78:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	53	93.0	109	2	S38496
2	50	87.7	96	2	S36060
3	50	87.7	106	2	S38495
4	50	87.7	108	2	S38498
5	50	87.7	108	2	S47184
6	50	87.7	109	2	S19663
7	50	87.7	110	2	S36272
8	50	87.7	115	2	S13726
9	50	87.7	127	2	S70444
10	49	86.0	233	2	S25748
11	45	78.9	233	2	S25741
12	44	77.2	2135	2	T14602
13	41	71.9	110	2	S19672
14	39	68.4	1729	2	T19396
15	39	68.4	2924	2	T18378
16	38	66.7	429	2	E64617
17	38	66.7	429	2	F71896
18	38	66.7	624	2	S44938
19	37	64.9	146	2	S02083
20	36	63.2	270	2	AG0916
21	36	63.2	452	1	WHHUF
22	36	63.2	547	2	D84942
23	36	63.2	1159	1	A44280
24	35	61.4	197	2	B83497
25	35	61.4	269	2	H84378
26	35	61.4	313	2	H72724
27	35	61.4	367	2	T24058
28	35	61.4	408	2	AC2188
29	35	61.4	561	2	JE0158

30	61.4	580	2	S72211	N-acetyl-beta-D-gl
31	61.4	868	2	AF3204	autotransporter pr
32	61.4	2228	2	T14029	variant-specific s
33	59.6	108	1	L3HUSH	Ig lambda chain V-
34	59.6	161	2	B44462	allophycocyanin be
35	59.6	189	2	D71978	hypothetical prote
36	59.6	189	2	G64527	hypothetical prote
37	59.6	218	2	G70438	hypothetical prote
38	59.6	284	2	T09452	vdIC protein - Hel
39	59.6	439	2	E86813	xylose isomerase (
40	59.6	446	2	T03267	probable phosphop
41	59.6	446	2	T02221	phosphoryruvate hy
42	59.6	448	1	WMBP82	gene 8 protein - p
43	59.6	464	2	F97807	sodium/pantothenat
44	59.6	529	2	A84260	hypothetical prote
45	59.6	561	2	T36084	hypothetical prote
46	59.6	586	2	T19381	hypothetical prote
47	59.6	657	2	S10001	MyD16 protein - m
48	59.6	695	2	E90152	hypothetical prote
49	59.6	812	2	A46417	NIPI protein - yea
50	57.9	117	2	E34792	Ig heavy chain pre
51	57.9	118	2	G34792	Ig heavy chain pre
52	57.9	145	2	S02025	hemoglobin beta ch
53	57.9	161	2	S33624	allophycocyanin be
54	57.9	178	2	B69944	hypothetical prote
55	57.9	235	2	G81219	hypothetical prote
56	57.9	235	2	C81991	hypothetical prote
57	57.9	260	2	C70675	hypothetical prote
58	57.9	268	2	H83571	rRNA (adenine-N6,N
59	57.9	270	2	F86897	hypothetical prote
60	57.9	289	1	A37209	thiosulfate sulfur
61	57.9	351	2	AC0302	conserved hypothet
62	57.9	374	2	T32286	hypothetical prote
63	57.9	387	2	S11985	serine proteinase
64	57.9	446	2	S16257	phosphoryruvate hy
65	57.9	448	1	WMBP89	gene 8 protein - p
66	57.9	453	1	WHRTF	phenylalanine 4-mo
67	57.9	453	2	S15758	phenylalanine 4-mo
68	57.9	470	2	T45962	hypothetical prote
69	57.9	558	2	S61604	probable membrane
70	57.9	571	2	T29751	hypothetical prote
71	57.9	592	2	T15600	hypothetical prote
72	57.9	709	2	S38241	hypothetical prote
73	57.9	751	2	T30154	hypothetical prote
74	57.9	764	2	AB1695	p60-related protei
75	57.9	1192	2	A71623	probable secreted
76	57.9	1327	2	D70759	probable ocsB prot
77	57.9	1785	2	S53976	probable membrane
78	56.1	102	2	G71137	hypothetical prote
79	56.1	114	2	B75117	hypothetical prote
80	56.1	114	2	S45541	hypothetical prote
81	56.1	148	2	T52550	probable transcrip
82	56.1	185	2	B82065	conserved hypothet
83	56.1	206	2	H64697	hypothetical prote
84	56.1	217	2	B71821	hypothetical prote
85	56.1	207	2	A42578	EF-hand calcium-bi
86	56.1	224	2	G90548	conserved hypothet
87	56.1	227	2	D84058	hypothetical prote
88	56.1	229	2	C43330	gene 7 protein - p
89	56.1	250	1	RGECF	transcription regu
90	56.1	250	2	AE0280	transcription regu
91	56.1	250	2	C90868	transcription reg
92	56.1	250	2	F85750	transcription reg
93	56.1	251	2	A12106	hypothetical prote
94	56.1	264	2	AE0662	transcription regu
95	56.1	266	2	F75081	hypothetical prote
96	56.1	293	2	F86070	probable enzyme 25
97	56.1	293	2	H91223	probable enzyme 1
98	56.1	332	2	G85187	hypothetical prote
99	56.1	342	2	T18994	hypothetical prote
100	56.1	343	2	AH0328	probable 3',5'-cyc
101	56.1	353	2	AB3650	iron(III)-transpor
102	56.1	395	2	B97728	hypothetical prote

```

103 32 56.1 407 2 E64208 oligopeptide trans
104 32 56.1 415 2 T0678 hypothetical prote
105 32 56.1 436 2 A69936 probable GTP bindi
106 32 56.1 443 2 AE3209 hypothetical prote
107 32 56.1 475 1 Z2BPL7 gene 2 protein - p
108 32 56.1 499 1 Z2BP22 gene 2 protein - p
109 32 56.1 501 2 T25317 hypothetical prote
110 32 56.1 515 2 T03070 hypothetical prote
111 32 56.1 553 2 B69167 glutamate-tRNA lig
112 32 56.1 559 2 J710949 egg-specific prote
113 32 56.1 561 2 S71189 Dwarfl protein - A
114 32 56.1 567 2 T06575 dwarf protein homo
115 32 56.1 569 2 H70076 hypothetical prote
116 32 56.1 574 2 D81372 probable outer-mem
117 32 56.1 594 2 B64018 hypothetical prote
118 32 56.1 595 1 YFYBVC phenylalanine-tRNA
119 32 56.1 629 2 A30168 homeotic protein L
120 32 56.1 635 2 S01164 homeotic protein L
121 32 56.1 662 2 G95929 probable methylcro
122 32 56.1 666 2 S29349 hypothetical prote
123 32 56.1 766 2 G69550 ATP-dependent RNA
124 32 56.1 926 2 A57072 disease resistance
125 32 56.1 996 2 D84561 probable AAA-type
126 32 56.1 1445 2 T14913 CAT8 protein - yea
127 32 56.1 1452 2 T22812 hypothetical prote
128 32 56.1 3973 2 B71612 hypothetical prote
129 31.5 55.3 367 1 TVCHFS transforming prote
130 31.5 55.3 370 2 A60089 transforming prote
131 31 54.4 86 2 S21595 Ig heavy chain v r
132 31 54.4 96 2 S04824 Ul snRNP 70K prote
133 31 54.4 98 2 S54856 Ig heavy chain v r
134 31 54.4 111 2 T13343 hypothetical prote
135 31 54.4 117 2 B34964 Ig heavy chain pre
136 31 54.4 145 2 E70113 probable transcrip
137 31 54.4 149 2 E71728 hypothetical prote
138 31 54.4 161 2 C82747 conserved hypotet
139 31 54.4 164 2 S74709 hypothetical prote
140 31 54.4 183 2 T28926 hypothetical prote
141 31 54.4 187 2 A65007 hypothetical prote
142 31 54.4 202 2 A82831 scol protein homol
143 31 54.4 205 2 F97608 scol protein homol
144 31 54.4 205 2 E72367 hypothetical prote
145 31 54.4 216 2 S00674 Ul snRNP 70K prote
146 31 54.4 225 2 F72642 probable snikmate
147 31 54.4 226 2 A91033 probable DNA trans
148 31 54.4 226 2 B85877 probable prophage
149 31 54.4 228 2 AE2173 hypothetical prote
150 31 54.4 235 2 B69436 hypothetical prote

```

ALIGNMENTS

```

RESULT 1
S38496
Ig lambda chain - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 21-Jan-2000
C:Accession: S38496
R:Marks, J.D.; Ouwehand, W.H.; Bye, J.M.; Finnern, R.; Gorick, B.D.; Voak, D.; Thorpe, S.
submitted to the EMBL Data Library, June 1993
A:Description: Human antibody fragments specific for human blood group antigens from a p
A:Reference number: S38488
A:Accession: S38496
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-109 <VAR>
A:Cross-references: EMBL:223031; NID:g414039; PIDN:CAA80566.1; PID:g414040
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-89/Domain: immunoglobulin homology <IMM>
Query Match 93.0%; Score 53; DB 2; Length 109;

```

```

Best Local Similarity 90.9%; Pred. No. 0.0039; Indels 0; Gaps 0;
Matches 10; Conservative 1; Mismatches 0;
QY 1 QGDSLKTYVAS 11
DB 23 QGDSLKTYVAS 33
RESULT 2
S36060
Ig lambda chain - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 22-Nov-1993 #sequence_revision 01-Dec-1995 #text_change 21-Jan-2000
C:Accession: S36060
R:Williams, S.C.
submitted to the EMBL Data Library, April 1993
A:Reference number: S36046
A:Accession: S36060
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-96 <WIL>
A:Cross-references: EMBL:22202; NID:g312325; PIDN:CAA80211.1; PID:g312326
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-89/Domain: immunoglobulin homology <IMM>
Query Match 87.7%; Score 50; DB 2; Length 96;
Best Local Similarity 81.8%; Pred. No. 0.013;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 QGDSLKTYVAS 11
DB 23 QGDSLRSYVAS 33
RESULT 3
S38495
Ig lambda chain - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 21-Jan-2000
C:Accession: S38495
R:Marks, J.D.; Ouwehand, W.H.; Bye, J.M.; Finnern, R.; Gorick, B.D.; Voak, D.; Thorpe, S.
submitted to the EMBL Data Library, June 1993
A:Description: Human antibody fragments specific for human blood group antigens from a p
A:Reference number: S38488
A:Accession: S38495
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-106 <MAR>
A:Cross-references: EMBL:223029; NID:g414037; PIDN:CAA80564.1; PID:g414038
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-89/Domain: immunoglobulin homology <IMM>
Query Match 87.7%; Score 50; DB 2; Length 106;
Best Local Similarity 81.8%; Pred. No. 0.015;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 QGDSLKTYVAS 11
DB 23 QGDSLRSYVAS 33
RESULT 4
S38498
Ig lambda chain - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 21-Jan-2000
C:Accession: S38498
R:Marks, J.D.; Ouwehand, W.H.; Bye, J.M.; Finnern, R.; Gorick, B.D.; Voak, D.; Thorpe, S.
submitted to the EMBL Data Library, June 1993
A:Description: Human antibody fragments specific for human blood group antigens from a p
A:Reference number: S38488

```

```
A:Accession: S38498
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-108 <VAR>
A:Cross-references: EMBL:223035; NID:g414043; PIDN:CAA80570.1; PID:g414044
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:14-88/Domain: immunoglobulin homology <IMM>

Query Match      87.7%; Score 50; DB 2; Length 108;
Best Local Similarity 81.8%; Pred. No. 0.015;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
   |||||:|||||
Db 22 QGDSLRSYYAS 32

RESULT 5
S47184
Ig lambda chain - human
C:Species: Homo sapiens (man)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 21-Jan-2000
C:Accession: S47184
R:McIntosh, R.S.; Tandon, N.; Metcalfe, R.A.; Weetman, A.P.
A:Description: Cloning and analysis of IgM anti-thyroglobulin autoantibodies from patient
submitted to the EMBL Data Library, June 1994
A:Reference number: S47181
A:Accession: S47184
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-108 <NCI>
A:Cross-references: EMBL:X79783; NID:g506426; PIDN:CAA56179.1; PID:g506427
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-89/Domain: immunoglobulin homology <IMM>

Query Match      87.7%; Score 50; DB 2; Length 108;
Best Local Similarity 81.8%; Pred. No. 0.015;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
   |||||:|||||
Db 23 QGDSLRSYYAS 33

RESULT 6
S19663
Ig lambda chain V region (clone alpha-BSA3) - human
C:Species: Homo sapiens (man)
C:Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 20-Jun-2000
C:Accession: S19663
R:Marks, J.D.; Hoogenboom, H.R.; Bonnert, T.P.; McCafferty, J.; Griffiths, A.D.; Winter,
J. Mol. Biol. 222, 581-597, 1991
A:Title: By-passing immunization. Human antibodies from V-gene libraries displayed on ph
A:Reference number: S19663; MUID:92085276; PMID:1748994
A:Accession: S19663
A:Molecule type: mRNA
A:Residues: 1-109 <VAR>
A:Cross-references: EMBL:X61640; NID:g29492; PIDN:CAA43821.1; PID:g1340166
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-89/Domain: immunoglobulin homology <IMM>

Query Match      87.7%; Score 50; DB 2; Length 109;
Best Local Similarity 81.8%; Pred. No. 0.015;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
   |||||:|||||
Db 23 QGDSLRSYYAS 33
```

RESULT 7

```
S36272
Ig lambda chain V region (clone alpha-THY-29) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 21-Jan-2000
C:Accession: S36272
R:Griffiths, A.D.; Malmqvist, M.; Marks, J.D.; Bye, J.M.; Embleton, M.J.; McCafferty, J.;
EMBO J. 12, 725-734, 1993
A:Title: Human anti-self antibodies with high specificity from phage display libraries.
A:Reference number: S36256; MUID:93178448; PMID:7679590
A:Accession: S36272
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-110 <GRI>
A:Cross-references: EMBL:Z18833; NID:g33419; PIDN:CAA79285.1; PID:g9339912
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-89/Domain: immunoglobulin homology <IMM>
```

```
Query Match      87.7%; Score 50; DB 2; Length 110;
Best Local Similarity 81.8%; Pred. No. 0.015;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 QGDSLKTYVAS 11
   |||||:|||||
Db 23 QGDSLRSYYAS 33
```

RESULT 8

```
S13726
Ig lambda chain V region - human
C:Species: Homo sapiens (man)
C:Date: 25-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 21-Jan-2000
C:Accession: S13726
R:Frippiat, J.P.; Chuchana, P.; Bernard, F.; Bulwela, L.; Lefranc, G.; Lefranc, M.P.
Nucleic Acids Res. 18, 7134, 1990
A:Title: First genomic sequence of a human Ig variable lambda gene belonging to subgroup
A:Reference number: S13726; MUID:91088295; PMID:2124677
A:Accession: S13726
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-115 <FRI>
A:Cross-references: EMBL:X56178; NID:g33404; PIDN:CAA39639.1; PID:g33405
C:Genetics:
A:Introns: 16/1
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:34-108/Domain: immunoglobulin homology <IMM>
```

```
Query Match      87.7%; Score 50; DB 2; Length 115;
Best Local Similarity 81.8%; Pred. No. 0.016;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 QGDSLKTYVAS 11
   |||||:|||||
Db 42 QGDSLRSYYAS 52
```

RESULT 9

```
S70444
Ig lambda chain precursor V region - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 21-Jan-2000
C:Accession: S70444; S70426
R:Cuisinier, A.M.; Fumoux, F.; Fougereau, M.; Tonnel, C.
Mol. Immunol. 29, 1363-1373, 1992
A:Title: IGM kappa/lambda EBV human B cell clone: an early step of differentiation of fet
A:Reference number: S70442; MUID:93024508; PMID:1383695
A:Accession: S70444
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-127 <CUI>
A:Experimental source: clone E29.1
```

R:Tonnelle, C.

submitted to the EMBL Data Library, May 1990

A:Reference number: S70426

A:Accession: S70426

A:Molecule type: mRNA

A:Residues: 1-90 <TON>

A:Cross-references: EMBL:X53070

A:Experimental source: cell line E29.1, clone VL 29-1

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:1-20/Domain: signal sequence #status predicted <SIG>

F:21-127/Product: Ig lambda chain V region (fragment) #status predicted <MAT>

F:148-108/Domain: immunoglobulin homology <IMM>

Query Match 87.7%; Score 50; DB 2; Length 127;

Best Local Similarity 81.8%; Pred. No. 0.018;

Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11

Db 42 QGDSLRSYYAS 52

RESULT 10

S25748

Ig lambda chain - human

C:Species: Homo sapiens (man)

C:Date: 22-Nov-1993 #sequence_revision 26-May-1995 #text_change 21-Jan-2000

R:Accession: S25748

R:Combiato, G.; Klobeck, H.G.

Eur. J. Immunol. 21, 1513-1522, 1991

A:Title: V(lambda) and J(lambda)-C(lambda) gene segments of the human immunoglobulin lam

A:Reference number: S16439; MUID:91257162; PMID:1904362

A:Accession: S25748

A:Status: preliminary; translation not shown

A:Molecule type: mRNA

A:Residues: 1-233 <COM>

A:Cross-references: EMBL:X57813; NID:g33725; PIDN:CAA40950.1; PID:g33726

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:148-216/Domain: immunoglobulin homology <IMM>

Query Match

Best Local Similarity 86.0%; Score 49; DB 2; Length 233;

Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11

Db 42 QGDSLRYAYAS 52

RESULT 11

S25741

Ig lambda chain - human

C:Species: Homo sapiens (man)

C:Date: 22-Nov-1993 #sequence_revision 26-May-1995 #text_change 21-Jan-2000

R:Accession: S25741

R:Combiato, G.; Klobeck, H.G.

Eur. J. Immunol. 21, 1513-1522, 1991

A:Title: V(lambda) and J(lambda)-C(lambda) gene segments of the human immunoglobulin lam

A:Reference number: S16439; MUID:91257162; PMID:1904362

A:Accession: S25741

A:Status: preliminary; translation not shown

A:Molecule type: mRNA

A:Residues: 1-233 <COM>

A:Cross-references: EMBL:X57805; NID:g33707; PIDN:CAA40943.1; PID:g33708

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:148-216/Domain: immunoglobulin homology <IMM>

Query Match

Best Local Similarity 78.9%; Score 45; DB 2; Length 233;

Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11

Db 42 QGDTIRSYVAS 52

RESULT 12

T14602

variant-specific surface protein - malaria parasite (Plasmodium falciparum) (fragments)

C:Species: Plasmodium falciparum

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jun-2000

R:Accession: T14602

R:Voss, T.S.; Felger, I.; Weiss, N.; Beck, H.P.

submitted to the EMBL Data Library, February 1998

A:Description: Identification of a conserved 5' flanking region of Plasmodium falciparum

A:Reference number: Z18158

A:Accession: T14602

A:Status: preliminary; translated from GB/EMBL/DBBJ

A:Molecule type: DNA

A:Residues: 1-2135 <VOS>

A:Cross-references: EMBL:AF050740; NID:g2944094; PID:g2944095; PIDN:AA05220.1

C:Genetics:

A:Gene: varph17

Query Match 77.2%; Score 44; DB 2; Length 2135;

Best Local Similarity 77.8%; Pred. No. 5.4;

Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYV 9

Db 160 EGDSTKTY 168

RESULT 13

S19672

Ig lambda chain V region (clone alpha-TEL14) - human

C:Species: Homo sapiens (man)

C:Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 20-Jun-2000

R:Accession: S19672

R:Marks, J.D.; Hoogenboom, H.R.; Bonnett, T.P.; McCafferty, J.; Griffiths, A.D.; Winter, J. Mol. Biol. 222, 581-597, 1991

A:Title: By-passing immunization. Human antibodies from V-gene libraries displayed on ph

A:Reference number: S19663; MUID:92085276; PMID:1748994

A:Accession: S19672

A:Molecule type: mRNA

A:Residues: 1-110 <MAR>

A:Cross-references: EMBL:X61644; NID:g37856; PIDN:CAA43825.1; PID:g1335384

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:15-89/Domain: immunoglobulin homology <IMM>

Query Match

Best Local Similarity 71.9%; Score 41; DB 2; Length 110;

Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11

Db 23 QGDSLRSYYAS 33

RESULT 14

T18396

erythrocyte membrane protein variant 2 - malaria parasite (Plasmodium falciparum) (fragme

C:Species: Plasmodium falciparum

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000

R:Accession: T18396

R:Baruch, D.I.; Pasloske, B.L.; Singh, H.B.; Bi, X.; Ma, X.C.; Feldman, M.; Taraschi, T.I.

Cell 82, 77-87, 1995

A:Title: Cloning the P. falciparum gene encoding PfEMP1, a malarial variant antigen and

A:Reference number: Z18925; MUID:95330812; PMID:7541722

A:Accession: T18396

A:Status: preliminary; translated from GB/EMBL/DBBJ

A:Molecule type: mRNA

A;Residues: 1-1729 <BAR>
A;Cross-references: EMBL:U27339; NID:g914920; PID:g914921; PIDN:AAA89134.1
C;Genetics:
A;Gene: EMP1
A;Note: var-2

Query Match 68.4%; Score 39; DB 2; Length 1729;
Best Local Similarity 66.7%; Pred. No. 42;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTTY 9
:||||:|
Db 151 EGDSEIKTHY 159

RESULT 15
TL8378
variant-specific surface protein 1 - malaria parasite (Plasmodium falciparum)
N;Alternate names: erythrocyte membrane binding protein 1 (EMP1)
C;Species: Plasmodium falciparum
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000
C;Accession: TL8378
R;Baruch, D.I.; Pastoske, B.L.; Singh, H.B.; Bi, X.; Ma, X.C.; Feldman, M.; Taraschi, T.
Cell 82, 77-87, 1995
A;Title: Cloning the P. falciparum gene encoding PfEMP1, a malarial variant antigen and
A;Reference number: Z18925; MUID:95330812; PMID:7541722
A;Accession: TL8378
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-2924 <BAR>
A;Cross-references: EMBL:U27338; NID:g914918; PID:g914919; PIDN:AAB60251.1
C;Genetics:
A;Gene: EMP1
A;Introns: 2476/3

Query Match 68.4%; Score 39; DB 2; Length 2924;
Best Local Similarity 66.7%; Pred. No. 73;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTTY 9
:||||:|
Db 151 EGDSEIKTHY 159

RESULT 16
E64617
hypothetical protein HP0781 - Helicobacter pylori (strain 26695)
C;Species: Helicobacter pylori
C;Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 08-Oct-1999
C;Accession: E64617
R;Tomb, J.F.; White, O.; Kervavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.;
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKen-
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.
Nature 388, 539-547, 1997
A;Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.
A;Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A;Reference number: A64520; MUID:97394467; PMID:9252185
A;Accession: E64617
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-429 <TOM>
A;Cross-references: GB:AE000590; GB:AE000511; NID:g2313907; PIDN:AAD07834.1; PID:g231391

Query Match 66.7%; Score 38; DB 2; Length 429;
Best Local Similarity 77.8%; Pred. No. 15;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 QGDSLKTTY 9
:||||:|
Db 74 QDESLKTTY 82

RESULT 17
E64617
hypothetical protein jhp0718 - Helicobacter pylori (strain J99)
C;Species: Helicobacter pylori
A;Variety: strain J99
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 08-Oct-1999
C;Accession: F71896
R;Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;
Ives, C.; Gibson, R.; Marberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; J.
Nature 397, 176-180, 1999
A;Title: Genomic sequence comparison of two unrelated isolates of the human gastric patho-
A;Reference number: A71800; MUID:99120557; PMID:9923682
A;Accession: F71896
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-429 <ARN>
A;Cross-references: GB:AE001503; GB:AE001439; NID:g4155275; PIDN:AAD06301.1; PID:g415528:
A;Experimental source: strain J99
C;Genetics:
A;Gene: jhp0718

Query Match 66.7%; Score 38; DB 2; Length 429;
Best Local Similarity 77.8%; Pred. No. 15;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 QGDSLKTTY 9
:||||:|
Db 74 QDESLKTTY 82

RESULT 18
S44938
nitrogen permease regulator NPR2 - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein YEL062w
C;Species: Saccharomyces cerevisiae
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 23-Mar-2001
C;Accession: S44938; S50527; S69008
R;Rousselet, G.; Simon, M.; Ripoché, P.; Buhler, J.M.
submitted to the EMBL Data Library, May 1994
A;Description: A second nitrogen permease regulator in Saccharomyces cerevisiae.
A;Reference number: S44938
A;Accession: S44938
A;Molecule type: DNA
A;Residues: 1-624 <ROU>
A;Cross-references: EMBL:X79105; NID:g485969; PID:g485970
R;Dietrich, F.S.
submitted to the EMBL Data Library, December 1994
A;Description: The sequence of S. cerevisiae cosmids 9669, 8334, 8199, and lambda clone 1
A;Reference number: S50434
A;Accession: S50527
A;Molecule type: DNA
A;Residues: 1-624 <DIE>
A;Cross-references: EMBL:U18795; NID:g603241; PID:g603256; MIPS:YEL062w
R;Rousselet, G.; Simon, M.; Ripoché, P.; Buhler, J.M.
FEBS Lett. 359, 215-219, 1995
A;Title: A second nitrogen permease regulator in Saccharomyces cerevisiae.
A;Reference number: S69008; MUID:95172238; PMID:7867803
A;Accession: S69008
A;Molecule type: DNA
A;Residues: 1-624 <ROW>
A;Cross-references: EMBL:X79105; NID:g485969; PIDN:CAA55721.1; PID:g485970
C;Genetics:
A;Gene: SGD:NPR2
A;Cross-references: SGD:S0000788; MIPS:YEL062w
A;Map position: 5L
C;Keywords: transcription regulation

Query Match 66.7%; Score 38; DB 2; Length 624;
Best Local Similarity 63.6%; Pred. No. 23;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 QGDSLKTTYAS 11
:||||:|
Db 456 QGQTLKTYES 466

RESULT 19

S02083
IG lambda chain V-IV region - human (tentative sequence) (fragments)
N:Alternate names: amyloid-fibril protein GIL
C:Species: Homo sapiens (man)
C>Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 31-Mar-2000
C:Accession: S02083
R:Fyke, E.M.; Sletten, K.; Rusby, G.; Cornwell III, G.G.
Biochem. J. 256, 973-980, 1988
A:Title: The primary structure of the variable region of an immunoglobulin IV light-chain
A:Reference number: S02083; MUID:89134210; PMID:3146981
A:Accession: S02083

A:Molecule type: protein
A:Residues: 1-70;71-72;73-75;76-131;132-146 <FYK>
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:14-88/Domain: immunoglobulin homology <IMM>

Query Match 64.9%; Score 37; DB 2; Length 146;
Best Local Similarity 63.6%; Pred. No. 7.9;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
|||||: : ||
DB 22 QGDSLRNFVAS 32

RESULT 20

AG0916
probable hydrolase STY3592 [imported] - Salmonella enterica subsp. enterica serovar Typh
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi

C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 28-Apr-2003
C:Accession: AG0916

R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001

A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AG0916

A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-270 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD07925.1; PID:g16504470; GSPDB:GN00176
A:Gene: STY3592
C:Superfamily: carboxymethylenebutenolidase

Query Match 63.2%; Score 36; DB 2; Length 270;
Best Local Similarity 60.0%; Pred. No. 24;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 QGDSLKTYVA 10
|||: |||
DB 38 QGDDMPAYYA 47

RESULT 21

WHHUF
phenylalanine 4-monooxygenase (EC 1.14.16.1) - human
N:Alternate names: phenylalanine 4-hydroxylase
C:Species: Homo sapiens (man)
C>Date: 28-Aug-1985 #sequence_revision 28-Aug-1985 #text_change 21-Jul-2000
C:Accession: A00508; S02687; I52416; I54346; S74142; I54257
R:Kwok, S.C.W.; Ledley, F.D.; DiLella, A.G.; Robson, K.J.H.; Woo, S.L.C.
Biochemistry 24, 556-561, 1985
A:Title: Nucleotide sequence of a full-length complementary DNA clone and amino acid seq
A:Reference number: A00508; MUID:85199778; PMID:2986678
A:Accession: A00508

A:Molecule type: mRNA

A:Residues: 1-452 <KWO>
A:Cross-references: GB:K03020; NID:gl89936; PIDN:AAAS0082.1; PID:gl89937
R:Cotton, R.G.H.; McAdam, W.; Jennings, I.; Morgan, F.J.
Biochem. J. 255, 193-196, 1988

A:Title: A monoclonal antibody to aromatic amino acid hydroxylases. Identification of the
A:Reference number: S02687; MUID:89061656; PMID:2461704

A:Accession: S02687

A:Molecule type: protein

A:Residues: 131-144 <COT>

R:Konecki, D.S.; Wang, Y.; Trefz, F.K.; Lichter-Konecki, U.; Woo, S.L.

Biochemistry 31, 8363-8368, 1992

A:Title: Structural characterization of the 5' regions of the human phenylalanine hydroxyl
A:Reference number: I52416; MUID:92399453; PMID:1326329

A:Accession: I52416

A>Status: preliminary; translated from GB/EMBL/DBBJ

A:Molecule type: DNA

A:Residues: 1-20 <KON>

A:Cross-references: GB:S44225; NID:9255493

R:Abadie, V.; Jaruzelska, J.; Lyonnet, S.; Millasseau, P.; Berthelon, M.; Rey, F.; Munnich

Hum. Mol. Genet. 2, 31-34, 1993

A:Title: Illegitimate transcription of the phenylalanine hydroxylase gene in lymphocytes

A:Reference number: I54346; MUID:93258345; PMID:8098245

A:Accession: I54346

A>Status: preliminary; translated from GB/EMBL/DBBJ

A:Molecule type: mRNA

A:Residues: 381-389; 'G', 391-405 <ABA>

A:Cross-references: GB:S61296; NID:G300410; PIDN:AAD13926.1; PID:g4261626

R:Experimental source: lymphocytes, mutant form

R:Kowlessur, D.; Citron, B.A.; Kaufman, S.

Arch. Biochem. Biophys. 333, 85-95, 1996

A:Title: Recombinant human phenylalanine hydroxylase: novel regulatory and structural pr
A:Reference number: S74142; MUID:96400381; PMID:8806757

A:Accession: S74142

A:Molecule type: protein

R:Rigel, A.; Dworniczak, B.; Kalaydjieva, L.; Horst, J.

Hum. Genet. 87, 739-743, 1991

A:Title: A frameshift mutation in exon 2 of the phenylalanine hydroxylase gene linked to

A:Reference number: I54257; MUID:92039642; PMID:1682235

A:Accession: I54257

A>Status: translated from GB/EMBL/DBBJ

A:Molecule type: DNA

A:Residues: 53-54; 'LRRNM' <ETG>

A:Cross-references: GB:S62592; NID:9238240; PIDN:AAB20205.1; PID:g238241

A:Note: mutant sequence found in patients with phenylketonuria

C:Genetics:

A:Gene: GDB:PAH

A:Cross-references: GDB:119470; OMIM:261600

A:Map position: 12q24.1-12q24.1

A:Note: a defect in this gene can cause phenylketonuria

C:Complex: homodimer

C:Function:

A:Description: catalyzes the 4'-hydroxylation of phenylalanine to tyrosine by tetrahydrot

A:Pathway: tyrosine biosynthesis; phenylalanine catabolism

C:Superfamily: phenylalanine 4-monooxygenase

C:Keywords: biotin; homodimer; iron; metalloprotein; monooxygenase; oxidoreductase; ph

F:16/Binding site: phosphate (Ser) (covalent) (by calmodulin-dependent kinase) #status pr

F:23/Binding site: phosphate (Ser) (covalent) (by cAMP-dependent kinase) #status predict

F:285,290,330/Binding site: iron (His, His, Glu) #status predicted

Query Match 63.2%; Score 36; DB 1; Length 452;
Best Local Similarity 70.0%; Pred. No. 41;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 QGDSLKTYVA 10
|||: |||

DB 336 QGDSIKAYCA 345

RESULT 22

DB4942
methionine-tRNA ligase (EC 6.1.1.10) [imported] - Buchnera sp. (strain APS)

N:Alternate names: methionyl-tRNA synthetase
C:Species: Buchnera sp.
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 17-Feb-2003
C:Accession: D84942
R:Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.
Nature 407, 81-86, 2000
A:Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. A
A:Reference number: A84930; MUID:20445173; PMID:10993077
A:Accession: D84942
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-547 <STO>
A:Cross-references: GB:AP000398; GSPDB:GN00144
A:Experimental source: strain APS
C:Genetics:
A:Gene: metG; BUI09
C:Superfamily: methionyl-tRNA synthetase
C:Keywords: ligase

Query Match 63.2%; Score 36; DB 2; Length 547;
Best Local Similarity 77.8%; Pred. No. 50;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 DSLKTYAS 11
DB 353 DSLRYYYAS 361
|||: ||||
|||: ||||

RESULT 23
A44280
inner layer protein VP1 - human rotavirus B (strain IDIR)
N:Alternate names: core protein VP1
C:Species: human rotavirus B
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 31-Dec-1993
C:Accession: A44280
R:Eiden, J.J.; Hirshon, C.
Virology 192, 154-160, 1993
A:Title: Sequence analysis of group B rotavirus gene 1 and definition of a rotavirus-spe
A:Reference number: A44280; MUID:93297100; PMID:8390749
A:Accession: A44280
A:Molecule type: Genomic RNA
A:Residues: 1-1159 <EID>
A:Cross-references: GB:M97203
C:Genetics:
A:Gene: 1
A:Map position: segment 1
C:Superfamily: rotavirus inner layer protein VP1
C:Keywords: core protein; glycoprotein
F:310,343,356,443,694,914/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 63.2%; Score 36; DB 1; Length 1159;
Best Local Similarity 70.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 GDSLKTYAS 11
DB 1039 GNTLKTLYAS 1048
||: |||||
||: |||||

RESULT 24
B83497
hypothetical protein PA1191 [imported] - Pseudomonas aeruginosa (strain PAO1)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: B83497
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Bu
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: B83497
A>Status: preliminary

A:Molecule type: DNA
A:Residues: 1-197 <STO>
A:Cross-references: GB:AE004549; GB:AE004091; NID:g9947110; PIDN:AAG04580.1; GSPDB:GN001
A:Experimental source: strain PAO1
C:Genetics:
A:Gene: PA1191

Query Match 61.4%; Score 35; DB 2; Length 197;
Best Local Similarity 66.7%; Pred. No. 27;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 QGDSLKTY 9
DB 101 QGDPLRDY 109
|||: ||
|||: ||

RESULT 25
H84378
hypothetical protein Vng2282c [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: H84378
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablor
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A:Title: Genome sequence of Halobacterium species NRC-1
A:Reference number: A84160; MUID:20504483; PMID:11016950
A:Accession: H84378
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-269 <STO>
A:Cross-references: GB:AE004437; NID:g10581693; PIDN:AAG20396.1; GSPDB:GN00138
C:Genetics:
A:Gene: VNG2282C

Query Match 61.4%; Score 35; DB 2; Length 269;
Best Local Similarity 60.0%; Pred. No. 37;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 GDSLKTYAS 11
DB 190 GDAIKSYLAS 199
||: |||
||: |||

Search completed: September 24, 2004, 01:54:23
Job time : 21.3962 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2005, 17:57:45 ; Search time 82.8438 Seconds
(without alignments)
51.354 Million cell updates/sec

Title: US-10-088-639a-2_COPY_23_33
Perfect score: 57
Sequence: 1 QGDSLKTYVAS 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_15Dec04:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001s:*
- 5: Geneseq2002s:*
- 6: Geneseq2003as:*
- 7: Geneseq2003bs:*
- 8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	57	100.0	249	4 AAB68087	Aab68087 An anti-a
2	54	94.7	11	8 ADS82613	Ads82613 Anti-IL-2
3	54	94.7	11	8 ADS82595	Ads82595 Anti-IL-2
4	54	94.7	104	2 AAR93164	Aar93164 Anti-rhes
5	54	94.7	109	8 ADS82590	Ads82590 Anti-IL-2
6	54	94.7	110	8 ADS82608	Ads82608 Anti-IL-2
7	54	94.7	238	2 AAR93165	Aar93165 Anti-rhes
8	54	94.7	242	8 ADI58098	Adi58098 Reg IV-sp
9	54	94.7	242	8 ADS82591	Ads82591 Anti-IL-2
10	54	94.7	245	8 ADS82609	Ads82609 Anti-IL-2
11	53	93.0	110	2 AAW15538	Aaw15538 Anti-TGF
12	53	93.0	253	5 ABP45519	Abp45519 Human Bly
13	53	93.0	253	5 ADG96346	Adg96346 Single ch
14	51	89.5	242	5 ABP46073	Abp46073 Human Bly
15	51	89.5	242	5 ADG96900	Adg96900 Single ch
16	51	89.5	243	5 ABP46091	Abp46091 Human Bly
17	51	89.5	243	7 ADG96918	Adg96918 Single ch
18	51	89.5	247	5 ABP46092	Abp46092 Human Bly
19	51	89.5	247	7 ADG96919	Adg96919 Single ch
20	51	89.5	248	5 ABP46077	Abp46077 Human Bly
21	51	89.5	248	7 ADG96904	Adg96904 Single ch
22	51	89.5	249	5 ABP45991	Abp45991 Human Bly
23	51	89.5	249	7 ADG96818	Adg96818 Single ch
24	51	89.5	250	5 ABP45979	Abp45979 Human Bly
25	51	89.5	250	7 ADG96806	Adg96806 Single ch

26	50	87.7	11	3 AAY95195	Aay95195 Anti-plat
27	50	87.7	11	3 AAY95216	Aay95216 Anti-plat
28	50	87.7	11	5 AAG80192	Aag80192 Human gp9
29	50	87.7	11	8 ADO58078	Ado58078 S9 cell d
30	50	87.7	96	7 ADJ80346	Adj80346 Vlamda g
31	50	87.7	101	2 AAR22572	Aar22572 Light cha
32	50	87.7	103	2 AAR80091	Aar80091 Human der
33	50	87.7	103	2 AAW95489	Aaw95489 Human-der
34	50	87.7	104	4 AAG93606	Aag93606 Human ant
35	50	87.7	104	5 ABR05054	Abro5054 Fibronect
36	50	87.7	104	6 ABO27413	Abo27413 Anti-Rh(D
37	50	87.7	106	4 AAU02531	Aau02531 Anti-adip
38	50	87.7	107	3 AAY95215	Aay95215 Anti-plat
39	50	87.7	107	3 AAY95194	Aay95194 Anti-plat
40	50	87.7	107	4 AAG62969	Aag62969 Amino aci
41	50	87.7	107	8 ADO36410	Ado36410 Intracell
42	50	87.7	108	3 AAY95185	Aay95185 Anti-plat
43	50	87.7	108	3 AAY95179	Aay95179 Anti-plat
44	50	87.7	108	3 AAY44616	Aay44616 Human ant
45	50	87.7	108	4 AAG62937	Aag62937 Amino aci

ALIGNMENTS

RESULT 1

AAB68087

ID AAB68087 standard; protein; 249 AA.

AC AAB68087;

XX 09-JUL-2001 (first entry)

DE An anti-alpha6beta4 integrin light chain linked to a heavy chain.

KW Gastrointestinal epithelial tumour cell; alpha6beta4 integrin;

KW tumour-associated antigen; metastatic disease; malignant disease;

KW tumour typing; tumour screening; tumour.

XX Synthetic.

OS Macaca fascicularis.

XX Key Location/Qualifiers

FT Region 23..33 /note= "Complementarity determining region (CDR) 1 of the light chain"

FT Region 49..55 /note= "Complementarity determining region (CDR) 2 of the light chain"

FT Region 88..98 /note= "Complementarity determining region (CDR) 3 of the light chain"

FT Peptide 110..127 /note= "linker"

FT Region 158..162 /note= "Complementarity determining region (CDR) 1 of the heavy chain"

FT Region 177..193 /note= "Complementarity determining region (CDR) 1 of the heavy chain"

FT Region 226..238 /note= "Complementarity determining region (CDR) 1 of the heavy chain"

XX WO200130854-A2.

XX 03-MAY-2001.

XX 26-OCT-2000; 2000WO-SE002082.

XX 28-OCT-1999; 99SE-00003895.

PA (ACTI-) ACTIVE BIOTECH AB.

XX Brodin TN, Karlstroem PJ, Ohlsson LG, Tordeson MJ, Kearney PP;
 PI Nilson BHK;
 XX WPI; 2001-308619/32.
 DR N-PSDB; AAP84797.
 XX
 PT Novel antibody for diagnosis, treatment of human metastatic and malignant
 PT diseases, has binding structure for target structure displayed on cell
 PT surface of human gastrointestinal epithelial tumor and normal cells.
 XX
 PS Claim 1; Page 55-56; 75pp; English.
 XX
 CC The present sequence represents a Monkey antibody light chain linked to a
 CC heavy chain. The antibody binds to a target structure displayed in and on
 CC the cell surface of human gastrointestinal epithelial tumour cells and in
 CC a subpopulation of normal human gastrointestinal epithelial cells. The
 CC target structure especially comprises alpha6beta4 integrin. This is a
 CC tumour-associated antigen. The antibody, and its fragments, are useful
 CC for treating conditions based on anti-angiogenic mechanism in humans.
 CC They are useful for treating human metastatic and malignant disease, for
 CC in vitro, in vivo diagnosis and prognosis of human malignant disease,
 CC comprising tumour typing, tumour screening, tumour diagnosis and
 CC prognosis and monitoring premalignant conditions. Quantitative in vivo
 CC diagnosis is carried out by determining the localization of antibody to
 CC tumour deposits in humans
 XX
 SQ Sequence 249 AA;
 Query Match 100.0%; Score 57; DB 4; Length 249;
 Best Local Similarity 100.0%; Pred. No. 0.034; 0; Indels 0; Gaps 0;
 Matches 11; Conservative 0; Mismatches 0;
 QY 1 QGDSLXYTYAS 11
 Db 23 QGDSLXYTYAS 33
 RESULT 2
 ADS82613
 ID ADS82613 standard; peptide; 11 AA.
 XX
 AC ADS82613;
 XX
 DT 16-DEC-2004 (first entry)
 XX
 DE Anti-IL-21R antibody 18A5 L1 CDR, SEQ ID 71.
 XX
 KW Immunosuppressive; Cytostatic; Antirheumatic; Antiarthritic;
 KW Antiinflammatory; Gastrointestinal; Antipsoriatic; Gene therapy;
 KW antibody; interleukin-21 receptor; interleukin-21; receptor; IL-21;
 KW IL-21R; autoimmune disorder; rheumatoid arthritis;
 KW inflammatory bowel disease; Crohn's disease; transplant rejection;
 KW psoriasis; hyperproliferative disorder; 18A5;
 KW complementarity determining region; CDR; light chain; L1 CDR.
 XX
 OS Homo sapiens.
 XX
 PF WO2004083249-A2.
 XX
 PN 30-SEP-2004.
 XX
 PD 12-MAR-2004; 2004WO-US007444.
 XX
 PF 14-MAR-2003; 2003US-0454336P.
 XX
 PR (AMHP) WYETH.
 PA (CAME-) CAMBRIDGE ANTIBODY TECHNOLOGY.
 XX
 PI Young DA, Whitters MJ, Valge-Archer V, Collins M, Williams AJ;
 PI Witek J;
 XX WPI; 2004-691025/67.

DR N-PSDB; ADS82622.
 XX
 PT New human antibodies that selectively bind to human interleukin-21
 PT receptor, useful for diagnosing, preventing or treating autoimmune
 PT disorders (e.g. rheumatoid arthritis) or hyperproliferative disorders.
 XX
 PS Claim 5; SEQ ID NO 71; 143pp; English.
 XX
 CC The present invention relates to human antibodies, or their antigen-
 CC binding fragments, that selectively bind to a human interleukin-21
 CC receptor (IL-21R). The antibodies of the invention are referred to as
 CC MUF, MUF-germline, MUI1, 18G4, 18A5, 19F5, CP5G2 and R18. The antibodies
 CC selectively bind the extracellular domain of human IL-21R, or inhibit the
 CC binding of IL-21 to an IL-21R. Pharmaceutical compositions comprising an
 CC antibody or fragment of the invention are useful for diagnosing,
 CC preventing or treating autoimmune disorders (e.g. rheumatoid arthritis,
 CC inflammatory bowel disease, Crohn's disease, transplant rejection or
 CC psoriasis) or hyperproliferative disorders. The present sequence is an
 CC anti-IL-21R antibody light chain complementarity determining region of
 CC the invention.
 XX
 SQ Sequence 11 AA;
 Query Match 94.7%; Score 54; DB 8; Length 11;
 Best Local Similarity 90.9%; Pred. No. 0.0043; 0; Indels 0; Gaps 0;
 Matches 10; Conservative 1; Mismatches 0;
 QY 1 QGDSLXYTYAS 11
 Db 1 QGDSLXYTYAS 11
 RESULT 3
 ADS82595
 ID ADS82595 standard; peptide; 11 AA.
 XX
 AC ADS82595;
 XX
 DT 16-DEC-2004 (first entry)
 XX
 DE Anti-IL-21R antibody 18G4 L1 CDR, SEQ ID 53.
 XX
 KW Immunosuppressive; Cytostatic; Antirheumatic; Antiarthritic;
 KW Antiinflammatory; Gastrointestinal; Antipsoriatic; Gene therapy;
 KW antibody; interleukin-21 receptor; interleukin-21; receptor; IL-21;
 KW IL-21R; autoimmune disorder; rheumatoid arthritis;
 KW inflammatory bowel disease; Crohn's disease; transplant rejection;
 KW psoriasis; hyperproliferative disorder; 18G4;
 KW complementarity determining region; CDR; light chain; L1 CDR.
 XX
 OS Homo sapiens.
 XX
 PN WO2004083249-A2.
 XX
 PD 30-SEP-2004.
 XX
 PF 12-MAR-2004; 2004WO-US007444.
 XX
 PN 14-MAR-2003; 2003US-0454336P.
 XX
 PD (AMHP) WYETH.
 PA (CAME-) CAMBRIDGE ANTIBODY TECHNOLOGY.
 XX
 PI Young DA, Whitters MJ, Valge-Archer V, Collins M, Williams AJ;
 PI Witek J;
 XX WPI; 2004-691025/67.
 DR N-PSDB; ADS82604.
 XX
 PT New human antibodies that selectively bind to human interleukin-21
 PT receptor, useful for diagnosing, preventing or treating autoimmune
 PT disorders (e.g. rheumatoid arthritis) or hyperproliferative disorders.

PS Claim 5; SEQ ID NO 53; 143pp; English.

XX The present invention relates to human antibodies, or their antigen-binding fragments, that selectively bind to a human interleukin-21 receptor (IL-21R). The antibodies of the invention are referred to as MUF, MUF-germline, MU11, 18G4, 18A5, 19F5, CP5G2 and R18. The antibodies selectively bind the extracellular domain of human IL-21R, or inhibit the binding of IL-21 to an IL-21R. Pharmaceutical compositions comprising an antibody or fragment of the invention are useful for diagnosing, preventing or treating autoimmune disorders (e.g. rheumatoid arthritis, inflammatory bowel disease, Crohn's disease, transplant rejection or psoriasis) or hyperproliferative disorders. The present sequence is an anti-IL-21R antibody light chain complementarity determining region of the invention.

XX SQ Sequence 11 AA;

Query Match 94.7%; Score 54; DB 8; Length 11;
Best Local Similarity 90.9%; Pred. No. 0.0043;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYTAS 11
| | | | | | | | | |
Db 1 QGDSLRTYYAS 11

RESULT 4

AA93164
ID AAR93164 standard; protein; 104 AA.

XX AC AAR93164;

XX 29-OCT-1996 (first entry)

DT Anti-rhesus D monoclonal antibody D7C2 light chain V region.

DE

XX Human monoclonal antibody; immunoglobulin isotype IgM; agglutination; rhesus positive; rhesus negative; haemolysis; lambda light chain; variable region; insect host cell; baculovirus; recombinant production.

KW

XX Homo sapiens.

OS

XX Key Location/Qualifiers

FT Region 23..33

FT /label= CDR1

FT /note= "complementarity determining region"

FT Region 49..55

FT /label= CDR2

FT /note= "complementarity determining region"

FT Region 88..93

FT /label= CDR3

FT /note= "complementarity determining region"

XX FR2724182-A1.

PN

XX 08-MAR-1996.

PD

XX 02-SEP-1994; 94FR-00010566.

PF

XX 02-SEP-1994; 94FR-00010566.

PR

XX (INSP) INST PASTEUR.

PA (PROT-) PROTEINE PERFORMANCE.

PA

XX Edelman L, Margartte C, Kaczorek M, Chaabihi H;

PI

XX WPI; 1996-162018/17.

DR

XX N-PSDB; AAT26869.

DR

XX Recombinant anti-rhesus D monoclonal antibody - expressed by baculovirus-transformed insect cells and useful for preventing haemolysis in new-born babies.

PT

XX

XX

PS Example 1; Page 30; 46pp; French.

XX The human monoclonal antibody D7C2, of isotype IgM, recognises a 30-32 kD polypeptide on the membrane of rhesus positive cells but not rhesus negative cells antibody agglutinates rhesus positive cells and also for preventing haemolysis in new-born rhesus positive babies. Recombinant IGM-D7C2 can be produced by insect cells which have been transformed by a baculoviral vector comprising a D7C2 expression cassette. The present sequence is that of the variable region of the IGM-D7C2 light chain

XX SQ Sequence 104 AA;

Query Match 94.7%; Score 54; DB 2; Length 104;
Best Local Similarity 90.9%; Pred. No. 0.048;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYTAS 11
| | | | | | | | | |
Db 23 QGDSLRTYYAS 33

RESULT 5

ADS82590
ID ADS82590 standard; protein; 109 AA.

XX AC ADS82590;

XX 16-DEC-2004 (first entry)

DT

XX Anti-IL-21R antibody 18G4 VL domain, SEQ ID 48.

DE

XX Immunosuppressive; Cytostatic; Antirheumatic; Antiarthritic;
Antinflammatory; Gastrointestinal; Antipsoriatic; Gene therapy;
antibody; interleukin-21 receptor; interleukin-21; receptor; IL-21;
IL-21R; autoimmune disorder; rheumatoid arthritis;
inflammatory bowel disease; Crohn's disease; transplant rejection;
psoriasis; hyperproliferative disorder; 18G4; variable region;
light chain.

KW

XX Homo sapiens.

OS

XX WO2004083249-A2.

PN

XX 30-SEP-2004.

PD

XX 12-MAR-2004; 2004WO-US007444.

PF

XX 14-MAR-2003; 2003US-0454336P.

PR

XX (AMHP) WYETH.

PA (CAMP-) CAMBRIDGE ANTIBODY TECHNOLOGY.

PA

XX Young DA, Whitters MJ, Valge-Archer V, Collins M, Williams AJ;
Witek J;

PI

XX WPI; 2004-691025/67.

DR

XX N-PSDB; ADS82599.

DR

XX New human antibodies that selectively bind to human interleukin-21 receptor, useful for diagnosing, preventing or treating autoimmune disorders (e.g. rheumatoid arthritis) or hyperproliferative disorders.

PT

XX Claim 1; SEQ ID NO 48; 143pp; English.

PS

XX The present invention relates to human antibodies, or their antigen-binding fragments, that selectively bind to a human interleukin-21 receptor (IL-21R). The antibodies of the invention are referred to as MUF, MUF-germline, MU11, 18G4, 18A5, 19F5, CP5G2 and R18. The antibodies selectively bind the extracellular domain of human IL-21R, or inhibit the binding of IL-21 to an IL-21R. Pharmaceutical compositions comprising an antibody or fragment of the invention are useful for diagnosing, preventing or treating autoimmune disorders (e.g. rheumatoid arthritis, inflammatory bowel disease, Crohn's disease, transplant rejection or psoriasis) or hyperproliferative disorders. The present sequence is an anti-IL-21R antibody light chain complementarity determining region of the invention.

XX SQ

CC inflammatory bowel disease, Crohn's disease, transplant rejection or
 CC psoriasis) or hyperproliferative disorders. The present sequence is an
 CC anti-IL-21R antibody light chain variable domain (VL) of the invention.
 XX
 SQ Sequence 109 AA;

Query Match 94.7%; Score 54; DB 8; Length 109;
 Best Local Similarity 90.9%; Pred. No. 0.05;
 Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
 |||||:|||||
 Db 23 QGDSLRTYYAS 33

RESULT 6
 ADS82608
 ID ADS82608 standard; protein; 110 AA.

XX AC ADS82608;
 XX DT 16-DEC-2004 (first entry)
 XX DE Anti-IL-21R antibody 18A5 VL domain, SEQ ID 66.
 XX KW Immunosuppressive; Cytostatic; Antirheumatic; Antiarthritic;
 KW Antinflammatory; Gastrointestinal; Antipsoriatic; Gene therapy;
 KW antibody; interleukin-21 receptor; interleukin-21; receptor; IL-21;
 KW IL-21R; autoimmune disorder; rheumatoid arthritis;
 KW inflammatory bowel disease; Crohn's disease; transplant rejection;
 KW psoriasis; hyperproliferative disorder; 18A5; variable region;
 KW light chain.

OS Homo sapiens.

XX WO2004083249-A2.

XX PD 30-SEP-2004.

XX PF 12-MAR-2004; 2004WO-US007444.

XX PR 14-MAR-2003; 2003US-0454336P.

XX PA (AMHP) WYETH.

XX PA (CAMP-) CAMBRIDGE ANTIBODY TECHNOLOGY.

XX PI Young DA, Whitters MJ, Valge-Archer V, Collins M, Williams AJ;
 PI Witek J;

XX DR WPI; 2004-691025/67.
 DR N-PSDB; ADS82617.

XX PT New human antibodies that selectively bind to human interleukin-21
 PT receptor, useful for diagnosing, preventing or treating autoimmune
 PT disorders (e.g. rheumatoid arthritis) or hyperproliferative disorders.

XX PS Claim 1; SEQ ID NO 66; 143pp; English.

XX CC The present invention relates to human antibodies, or their antigen-
 CC binding fragments, that selectively bind to a human interleukin-21
 CC receptor (IL-21R). The antibodies of the invention are referred to as
 CC MAF, MAF-germline, MUI1, 18A5, 18A5, 19F5, CFSG2 and R18. The antibodies
 CC selectively bind the extracellular domain of human IL-21R, or inhibit the
 CC binding of IL-21 to an IL-21R. Pharmaceutical compositions comprising an
 CC antibody or fragment of the invention are useful for diagnosing,
 CC preventing or treating autoimmune disorders (e.g. rheumatoid arthritis,
 CC inflammatory bowel disease, Crohn's disease, transplant rejection or
 CC psoriasis) or hyperproliferative disorders. The present sequence is an
 CC anti-IL-21R antibody light chain variable domain (VL) of the invention.

XX SQ Sequence 110 AA;

Query Match 94.7%; Score 54; DB 8; Length 110;

Best Local Similarity 90.9%; Pred. No. 0.051;
 Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
 |||||:|||||
 Db 23 QGDSLRTYYAS 33

RESULT 7

AAR93165

ID AAR93165 standard; protein; 238 AA.

XX AC AAR93165;

XX DT 30-OCT-1996 (first entry)

XX DE Anti-rhesus D recombinant antibody D7C2 light chain.

XX KW Human monoclonal antibody; immunoglobulin isotype IgM; agglutination;
 KW rhesus positive; rhesus negative; haemolysis; lambda light chain;
 KW variable region; insect host cell; baculovirus; recombinant production.

XX OS Homo sapiens.

XX OS Synthetic.

XX FH Key Location/Qualifiers

XX FT Peptide 1..19

XX FT /label= signal peptide

XX FT /note= "encoded by synthetic linker"

XX FT Protein 20..238

XX FT /label= light chain

XX FT /note= "human lambda light chain constant region and the
 FT variable region from anti-rhesus D antibody D7C2"

XX PN FR2724182-A1.

XX PD 08-MAR-1996.

XX PF 02-SEP-1994; 94FR-00010566.

XX PR 02-SEP-1994; 94FR-00010566.

XX PA (INSP) INST PASTEUR.

XX PA (PROT-) PROTEINE PERFORMANCE.

XX PI Edelman L, Margaritte C, Kaczorek M, Chaabihi H;

XX DR WPI; 1996-162018/17.

XX DR N-PSDB; AAT26888.

XX PT Recombinant anti-rhesus D monoclonal antibody - expressed by baculovirus-
 PT transformed insect cells and useful for preventing haemolysis in new-born
 PT babies.

XX PS Example 2; Page 33-34; 46pp; French.

XX CC The human monoclonal antibody D7C2, of isotype IgM, recognises a 30-32 kD
 CC polypeptide on the membrane of rhesus positive red blood cells. The
 CC antibody agglutinates rhesus positive cells but not rhesus negative cells
 CC and is useful diagnostically and also for preventing haemolysis in new-
 CC born rhesus positive babies. Recombinant IgM-D7C2 can be produced by
 CC insect cells which have been transformed by a baculoviral vector
 CC comprising a D7C2 expression cassette. The present sequence is that of a
 CC recombinant IgM-D7C2 lambda light chain fused downstream of a signal
 CC peptide

XX SQ Sequence 238 AA;

Query Match 94.7%; Score 54; DB 2; Length 238;
 Best Local Similarity 90.9%; Pred. No. 0.12;
 Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11

```

Db      42 QGDSLRTYYAS 52
||||:|||||
RESULT 8
ADIS8098
ID ADIS8098 standard; protein; 242 AA.
XX
XX AC ADIS8098;
XX
XX DT 22-APR-2004 (first entry)
XX
XX DE Reg IV-specific single chain antibody fragment (scFv) #57.
XX
XX KW antibody; regeneration IV; Reg IV; single chain antibody fragment; scFv;
XX inflammatory bowel disorder; ulcerative colitis; Crohn's disease;
XX KW diabetes; non-insulin dependent diabetes; insulin dependent diabetes;
XX cancer; human.
XX
XX OS Homo sapiens.
XX
XX PN WO2004003144-A2.
XX
XX PD 08-JAN-2004.
XX
XX PF 26-JUN-2003; 2003WO-US019908.
XX
XX PR 01-JUL-2002; 2002US-0392382P.
XX
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX PI Rosen CA;
XX
XX DR WPI: 2004-071976/07.
XX N-PSDB; ADIS8164.
XX
XX PT Novel antibody, useful for treating, preventing or ameliorating
XX inflammatory bowel disorder, cancer of the gastrointestinal tract or
XX PT diabetes (non-insulin dependent diabetes or insulin dependent diabetes).
XX
XX PS Claim 2; SEQ ID NO 58; 324pp; English.
XX
XX CC The invention comprises an antibody that specifically binds a
XX CC regeneration IV (Reg IV) protein. The invention specifically comprises
XX CC the amino acid and coding sequences of single chain antibody fragments
XX CC (scFv's) that bind Reg IV protein. The antibody of the invention is
XX CC useful for treating, preventing and ameliorating: inflammatory bowel
XX CC disorders (e.g. ulcerative colitis or Crohn's disease), diabetes (e.g.
XX CC non-insulin dependent diabetes or insulin dependent diabetes), and cancer
XX CC of the gastrointestinal tract. The antibody of the invention is also
XX CC useful for detecting the expression of a Reg IV protein. The present
XX CC amino acid sequence represents an scFv of the invention.
XX
XX SQ Sequence 242 AA;
Query Match 94.7%; Score 54; DB 8; Length 242;
Best Local Similarity 90.9%; Pred. No. 0.12;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 QGDSLRTYYAS 11
Db 156 QGDSLRTYYAS 166
RESULT 9
ADIS82591
ID ADS82591 standard; protein; 242 AA.
XX
XX AC ADS82591;
XX
XX DT 16-DEC-2004 (first entry)
XX
XX DE Anti-IL-21R antibody 18G4 scFv, SEQ ID 49.
XX
XX KW Immunosuppressive; Cytostatic; Antirheumatic; Antiarthritic;
XX Antinflammatory; Gastrointestinal; Antipsoriatic; Gene therapy;
XX KW antibody; interleukin-21 receptor; interleukin-21; receptor; IL-21;
XX IL-21R; autoimmune disorder; rheumatoid arthritis;
XX KW inflammatory bowel disease; Crohn's disease; transplant rejection;
XX KW psoriasis; hyperproliferative disorder; 18G4; scFv.
XX
XX OS Homo sapiens.
XX
XX PH Key Location/Qualifiers
XX FT Region 31..35
XX FT /label= H1_CDR
XX FT Region 50..66
XX FT /label= H2_CDR
XX FT Region 99..105
XX FT /label= H3_CDR
XX FT Region 156..166
XX FT /label= L1_CDR
XX FT Region 182..188
XX FT /label= L2_CDR
XX FT Region 221..231
XX FT /label= L3_CDR
XX
XX PN WO2004083249-A2.
XX
XX PD 30-SEP-2004.
XX
XX PF 12-MAR-2004; 2004WO-US007444.
XX
XX PR 14-MAR-2003; 2003US-0454336P.
XX
XX PA (AMHP ) WYETH.
XX (CAMP-) CAMBRIDGE ANTIBODY TECHNOLOGY.
XX
XX PI Young DA, Whitters MJ, Valge-Archer V, Collins M, Williams AJ;
XX Wittek J;
XX
XX DR WPI: 2004-691025/67.
XX N-PSDB; ADS82600.
XX
XX PT New human antibodies that selectively bind to human interleukin-21
XX PT receptor, useful for diagnosing, preventing or treating autoimmune
XX PT disorders (e.g. rheumatoid arthritis) or hyperproliferative disorders.
XX
XX PS Claim 1; SEQ ID NO 49; 143pp; English.
XX
XX CC The present invention relates to human antibodies, or their antigen-
XX CC binding fragments, that selectively bind to a human interleukin-21
XX CC receptor (IL-21R). The antibodies of the invention are referred to as
XX CC MUF, MUF-germline, MUI1, 18G4, 18A5, 19P5, CP5G2 and R18. The antibodies
XX CC selectively bind the extracellular domain of human IL-21R, or inhibit the
XX CC binding of IL-21 to an IL-21R. Pharmaceutical compositions comprising an
XX CC antibody or fragment of the invention are useful for diagnosing,
XX CC preventing or treating autoimmune disorders (e.g. rheumatoid arthritis,
XX CC inflammatory bowel disease, Crohn's disease, transplant rejection or
XX CC psoriasis) or hyperproliferative disorders. The present sequence is an
XX CC anti-IL-21R antibody scFv domain of the invention.
XX
XX SQ Sequence 242 AA;
Query Match 94.7%; Score 54; DB 8; Length 242;
Best Local Similarity 90.9%; Pred. No. 0.12;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 QGDSLRTYYAS 11
Db 156 QGDSLRTYYAS 166
RESULT 10
ADS82609
ID ADS82609 standard; protein; 245 AA.

```

XX ADS82609;
 AC
 XX
 DT 16-DEC-2004 (first entry)
 XX
 DE DE
 DE Anti-IL-21R antibody 18A5 scFv, SEQ ID 67.
 XX
 XX Immunosuppressive; Cytostatic; Antirheumatic; Antiarthritic;
 KW Antinflammatory; Gastrointestinal; Antipsoriatic; Gene therapy;
 KW antibody; Interleukin-21 receptor; Interleukin-21; receptor; IL-21;
 KW IL-21R; autoimmune disorder; rheumatoid arthritis;
 KW inflammatory bowel disease; Crohn's disease; transplant rejection;
 KW psoriasis; hyperproliferative disorder; 18A5; scFv.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH Region 31..36
 FT /label= H1_CDR
 FT Region 51..66
 FT /label= H2_CDR
 FT Region 99..107
 FT /label= H3_CDR
 FT Region 158..168
 FT /label= L1_CDR
 FT Region 184..190
 FT /label= L2_CDR
 FT Region 223..234
 FT /label= L3_CDR
 XX
 XX WO2004083249-A2.
 PN
 XX
 XX 30-SEP-2004.
 PD
 XX
 XX 12-MAR-2004; 2004WO-US007444.
 PF
 XX
 XX 14-MAR-2003; 2003US-0454336P.
 PR
 XX
 XX (AMHP) WYETH.
 PA (CAMB-) CAMBRIDGE ANTIBODY TECHNOLOGY.
 XX
 XX Young DA, Whitters MJ, Valge-Archer V, Collins M, Williams AJ;
 PI Witek J;
 PI
 XX WPI; 2004-691025/67.
 DR
 DR N-PSDB; ADS82618.
 XX
 XX New human antibodies that selectively bind to human interleukin-21
 PT receptor, useful for diagnosing, preventing or treating autoimmune
 PT disorders (e.g. rheumatoid arthritis) or hyperproliferative disorders.
 XX
 PS Claim 1; SEQ ID NO 67; 143pp; English.
 XX
 CC The present invention relates to human antibodies, or their antigen-
 CC binding fragments, that selectively bind to a human interleukin-21
 CC receptor (IL-21R). The antibodies of the invention are referred to as
 CC MUF, MUF-germline, MUI1, 18G4, 18A5, 19F5, CP5G2 and R18. The antibodies
 CC selectively bind the extracellular domain of human IL-21R, or inhibit the
 CC binding of IL-21 to an IL-21R. Pharmaceutical compositions comprising an
 CC antibody or fragment of the invention are useful for diagnosing,
 CC preventing or treating autoimmune disorders (e.g. rheumatoid arthritis,
 CC inflammatory bowel disease, Crohn's disease, transplant rejection or
 CC psoriasis) or hyperproliferative disorders. The present sequence is an
 CC anti-IL-21R antibody scFv domain of the invention.
 XX
 XX Sequence 245 AA;
 SQ
 Query Match 94.7%; Score 54; DB 8; Length 245;
 Best Local Similarity 90.9%; Pred. No. 0.12;
 Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QGDSLKTYVAS 11
 |||||:|||||

Db 158 QGDSLRTYYAS 168
 RESULT 11
 AAW15538
 ID AAW15538 standard; protein; 110 AA.
 XX
 AC AAW15538;
 XX
 DT 27-NOV-1997 (first entry)
 XX
 DE Anti-TGF beta-1 scFv antibody 10A6 VL domain.
 XX
 KW Transforming growth factor beta-1; TGF-beta-1; human;
 KW antibody engineering; scFv; phage display; lung fibrosis;
 KW arterial injury; proliferative retinopathy; retinal detachment;
 KW adult respiratory distress syndrome; liver cirrhosis;
 KW post myocardial infarction; post-angioplasty restenosis; scleroderma;
 KW vascular disease; cataract; glaucoma; scarring; glomerulonephritis;
 KW osteoporosis; immune disease; inflammation; rheumatoid arthritis;
 KW macrophage deficiency disease; macrophage pathogen infection; therapy.
 XX
 OS Homo sapiens.
 XX
 XX GB2305921-A.
 PN
 XX
 XX 23-APR-1997.
 PD
 XX
 XX 07-OCT-1996; 96GB-00020920.
 PF
 XX
 XX 06-OCT-1995; 95GB-00020486.
 PR
 XX 19-JAN-1996; 96GB-00001081.
 PR
 XX
 XX (CAMB-) CAMBRIDGE ANTIBODY TECHNOLOGY.
 PA
 XX Thompson JE, Vaughan TJ, Williams AJ, Green JA, Jackson RH;
 PI Bacon L, Johnson KS, Wilton AJ, Tempest PR, Pope AR;
 PI
 XX WPI; 1997-215360/20.
 DR
 DR N-PSDB; AAT60384.
 XX
 XX Agent contg. antigen-binding domain of human antibody to transforming
 PT growth factor beta 1 or 2 - and nucleic acid encoding it, used to
 PT neutralise effects of TGF, e.g. for control of fibrosis, immune and
 PT inflammatory disease.
 XX
 XX Claim 16; Fig 1b(iii); 184pp; English.
 PS
 CC This polypeptide sequence comprises the VL domain of human scFv antibody
 CC 10A6, which is specific for transforming growth factor (TGF) beta-1. It
 CC is encoded by a gene (AAT60384) isolated from a peripheral blood
 CC lymphocyte library. The antigen-binding domains of human antibodies (see
 CC AAW15522-40) to TGF beta-1 and/or beta-2 can be used to counter the
 CC adverse effects of TGF beta, such as (i) promotion of fibrosis (in
 CC dermal, ocular or keloid scarring, lung fibrosis, arterial injury,
 CC proliferative retinopathy, retinal detachment, adult respiratory distress
 CC syndrome, liver cirrhosis, post myocardial infarction, post-angioplasty
 CC restenosis, scleroderma, vascular disorders, cataract, glaucoma, or esp.
 CC neural scarring and glomerulonephritis, also (not claimed) osteoporosis),
 CC or (ii) immune and inflammatory diseases (e.g. rheumatoid arthritis,
 CC macrophage deficiency diseases or macrophage pathogen infection). Nucleic
 CC acids encoding human antibody VH and VL can be used for prodn. of
 CC recombinant antigen-binding domains. These are highly specific, have low
 CC dissociation constants (pref. less than 5 nM) and low IC50s for
 CC neutralisation
 XX
 XX Sequence 110 AA;
 SQ
 Query Match 93.0%; Score 53; DB 2; Length 110;
 Best Local Similarity 90.9%; Pred. No. 0.078;
 Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QGDSLKTYVAS 11

```
Db          23 QGDSLKSYVAS 33
|||||:||||
RESULT 12
ABP45519
ID ABP45519 standard; protein; 253 AA.
AC ABP45519;
XX
DT 19-AUG-2002 (first entry)
XX
DE Human BlyS binding scFv SEQ ID 1530.
XX
KW BlyS; B lymphocyte stimulator; TNF superfamily; human; cytostatic;
KW tumour necrosis factor; B cell proliferation; B cell differentiation;
KW immunosuppressive; immunostimulant; immunomodulatory; antirheumatic;
KW antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;
KW systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;
KW common variable immunodeficiency; acquired immunodeficiency syndrome.
XX
OS Homo sapiens.
XX
PN WO200202641-A1.
XX
PD 10-JAN-2002.
XX
PP 15-JUN-2001; 2001WO-US019110.
XX
PR 16-JUN-2000; 2000US-0212210P.
XX
PR 17-OCT-2000; 2000US-0240816P.
XX
PR 16-MAR-2001; 2001US-0276248P.
XX
PR 21-MAR-2001; 2001US-0277379P.
XX
PR 25-MAY-2001; 2001US-0293499P.
XX
(HUMA-) HUMAN GENOME SCI INC.
PA
PA (CAMB-) CAMBRIDGE ANTIBODY TECHNOLOGY.
PI
PI Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;
XX
DR WPI; 2002-114799/15.
XX
Antibodies against B Lymphocyte Stimulating polypeptides, useful for the
diagnosis and treatment of cancers and immune disorders.
XX
Claim 1; Page 2228-2229; 3148pp; English.
XX
This invention describes novel antibodies that immunospecifically bind to
B Lymphocyte Stimulator (BlyS) polypeptides. BlyS is a member of the
tumour necrosis factor (TNF) super family and induces B cell
proliferation and differentiation. The antibodies of the invention have
cytostatic, immunosuppressive, immunostimulant, immunomodulatory,
antirheumatic and antiAIDS activity and can be used in vaccines to
inhibit the expression and activity of BlyS. The antibodies bind to BlyS
and so may be used to detect and quantitate the presence of BlyS in
biological samples and may be used in this way to diagnose disease
associated with aberrant expression of BlyS. They may also be
administered to treat diseases associated with aberrant BlyS expression
and activity such as cancer, immune, and autoimmune disorders and
diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis,
immunodeficiency (e.g. common variable immunodeficiency (CVID) and
acquired immunodeficiency syndrome (AIDS)). ABP43990-ABP47228 represent
the antibodies and fragments of the antibodies described in the method of
the invention
XX
SQ Sequence 253 AA;
Query Match 93.0%; Score 53; DB 5; Length 253;
Best Local Similarity 90.9%; Pred. No. 0.19;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 QGDSLKTYVAS 11
|||||:||||

Db          166 QGDSLKSYVAS 176
RESULT 13
ADG96346
ID ADG96346 standard; protein; 253 AA.
AC ADG96346;
XX
DT 11-MAR-2004 (first entry)
XX
DE Single chain antibody that immunospecifically binds BlyS SeqID 1530.
XX
KW antibody; B lymphocyte stimulator; BlyS; tumour necrosis factor;
KW B cell proliferation; differentiation; scFv; myasthenia gravis;
KW multiple sclerosis; asthma; rheumatoid arthritis; AIDS; leukaemia;
KW carcinoma; lymphoma; antirheumatic; antiarthritic; neuroprotective;
KW antiinflammatory; antiasthmatic; antiallergic; cytostatic.
XX
OS Unidentified.
XX
PN WO2003055979-A2.
XX
PD 10-JUL-2003.
XX
PP 14-NOV-2002; 2002WO-US036496.
XX
PR 16-NOV-2001; 2001US-0331469P.
XX
PR 19-DEC-2001; 2001US-0340817P.
XX
(HUMA-) HUMAN GENOME SCI INC.
PA
PA Ruben SM, Barash SC, Choi GH, Vaughan TJ, Hilbert D;
XX
DR WPI; 2003-505530/47.
XX
Novel antibody that immunospecifically binds to a B lymphocyte stimulator
(BlyS), useful for detecting and treating diseases or disorders e.g.
rheumatoid arthritis, asthma and leukemia.
XX
Example 1; SEQ ID NO 1530; 394pp; English.
XX
This invention relates to novel antibodies that immunospecifically bind
to B lymphocyte stimulator (BlyS). The BlyS gene has been mapped to
chromosome 13q34 and encodes a protein that is a member of the tumour
necrosis factor superfamily and induces both in vivo and in vitro B cell
proliferation and differentiation. Specifically, it refers to single
chain antibody molecules (scFvs) derived, preferably, from the variable
heavy CDR3 region that immunospecifically bind to a polypeptide, or
fragment thereof, of either human, murine, rat or monkey BlyS. The
present invention refers to the use of such antibodies in various methods
for the detection, diagnosis and prognosis of diseases related to the
aberrant expression or inappropriate function of BlyS or its receptor. As
such, these compositions are useful for identifying immune disorders
including myasthenia gravis and multiple sclerosis, inflammatory
disorders e.g. asthma and rheumatoid arthritis, infectious diseases such
as AIDS and proliferative disorders including leukaemia, carcinoma and
lymphoma. Accordingly, they can be described as exhibiting various
activities such as antirheumatic, antiarthritic, neuroprotective,
antiinflammatory, antiasthmatic, antiallergic and cytostatic. This
polypeptide sequence is a single chain antibody that binds BlyS of the
invention. NOTE: The sequence data for this patent did not form part of
the printed specification, but was obtained in electronic format
directly from WIPO at ftp.wipo.int/pub/published pct_sequences.
XX
SQ Sequence 253 AA;
Query Match 93.0%; Score 53; DB 7; Length 253;
Best Local Similarity 90.9%; Pred. No. 0.19;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 QGDSLKTYVAS 11
|||||:||||
```

Db 166 QGDSLKSYVAS 176

RESULT 14
ABP46073
ID ABP46073 standard; protein; 242 AA.
XX
AC ABP46073;
XX
DT 19-AUG-2002 (first entry)
XX
DE Human BlyS binding scFv SEQ ID 2084.
XX
KW BlyS; B lymphocyte stimulator; TNF superfamily; human; cytostatic;
KW tumour necrosis factor; B cell proliferation; B cell differentiation;
KW immunosuppressive; immunostimulant; immunomodulatory; antirheumatic;
KW antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;
KW systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;
KW common variable immunodeficiency; acquired immunodeficiency syndrome.
XX
OS Homo sapiens.
XX
XX WO200202641-A1.
XX
XX 10-JAN-2002.
XX
XX 15-JUN-2001; 2001WO-US019110.
XX
XX 16-JUN-2000; 2000US-0212210P.
PR 17-OCT-2000; 2000US-0240816P.
PR 16-MAR-2001; 2001US-0276248P.
PR 21-MAR-2001; 2001US-0277379P.
PR 25-MAY-2001; 2001US-0293499P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA (CAMB-) CAMBRIDGE ANTIBODY TECHNOLOGY.
XX
XX Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;
XX WPI; 2002-114799/15.
DR
XX
XX Antibodies against B Lymphocyte Stimulating polypeptides, useful for the
XX diagnosis and treatment of cancers and immune disorders.
XX
XX Claim 1; Page 2883-2885; 3148pp; English.
PS
XX This invention describes novel antibodies that immunospecifically bind to
XX B Lymphocyte Stimulator (BlyS) polypeptides. BlyS is a member of the
XX tumour necrosis factor (TNF) super family and induces B cell
XX proliferation and differentiation. The antibodies of the invention have
XX cytostatic, immunosuppressive, immunostimulant, immunomodulatory,
XX antirheumatic and antiAIDS activity and can be used in vaccines to
XX inhibit the expression and activity of BlyS. The antibodies bind to BlyS
XX and so may be used to detect and quantitate the presence of BlyS in
XX biological samples and may be used in this way to diagnose disease
XX associated with aberrant expression of BlyS. They may also be
XX administered to treat diseases associated with aberrant BlyS expression
XX and activity such as cancer, immune, and autoimmune disorders and
XX diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis,
XX immunodeficiency (e.g. common variable immunodeficiency (CVID) and
XX acquired immunodeficiency syndrome (AIDS)). ABP43990-ABP47228 represent
XX the antibodies and fragments of the antibodies described in the method of
XX the invention

Query Match 89.5%; Score 51; DB 5; Length 242;
Best Local Similarity 81.8%; Pred. No. 0.43;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
Db 156 QGDSLRITYAN 166

RESULT 15
ADG96900
ID ADG96900 standard; protein; 242 AA.
XX
AC ADG96900;
XX
DT 11-MAR-2004 (first entry)
XX
DE Single chain antibody that immunospecifically binds BlyS SeqID 2084.
XX
KW antibody; B lymphocyte stimulator; BlyS; tumour necrosis factor;
KW B cell proliferation; differentiation; scFv; myasthenia gravis;
KW multiple sclerosis; asthma; rheumatoid arthritis; AIDS; leukaemia;
KW carcinoma; lymphoma; antirheumatic; antiarthritic; neuroprotective;
KW antinflammatory; antiasthmatic; anti-allergic; cytostatic.
XX
OS Unidentified.
XX
XX WO2003055979-A2.
PN
XX 10-JUL-2003.
PD
XX 14-NOV-2002; 2002WO-US036496.
PF
XX 16-NOV-2001; 2001US-0331469P.
PR 19-DEC-2001; 2001US-0340817P.
PR
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX Ruben SM, Barash SC, Choi GH, Vaughan TJ, Hilbert D;
PI WPI; 2003-505530/47.
XX
XX Novel antibody that immunospecifically binds to a B lymphocyte stimulator
XX (BlyS), useful for detecting and treating diseases or disorders e.g.
XX rheumatoid arthritis, asthma and leukemia.
PS Example 1; SEQ ID NO 2084; 394pp; English.
XX
XX This invention relates to novel antibodies that immunospecifically bind
XX to B lymphocyte stimulator (BlyS). The BlyS gene has been mapped to
XX chromosome 13q34 and encodes a protein that is a member of the tumour
XX necrosis factor superfamily and induces both in vivo and in vitro B cell
XX proliferation and differentiation. Specifically, it refers to single
XX chain antibody molecules (scFvs) derived, preferably, from the variable
XX heavy CDR3 region that immunospecifically bind to a polypeptide, or
XX fragment thereof, of either human, murine, rat or monkey BlyS. The
XX present invention refers to the use of such antibodies in various methods
XX for the detection, diagnosis and prognosis of diseases related to the
XX aberrant expression or inappropriate function of BlyS or its receptor. As
XX such, these compositions are useful for identifying immune disorders
XX including myasthenia gravis and multiple sclerosis, inflammatory
XX disorders e.g. asthma and rheumatoid arthritis, infectious diseases such
XX as AIDS and proliferative disorders including leukaemia, carcinoma and
XX lymphoma. Accordingly, they can be described as exhibiting various
XX activities such as antirheumatic, antiarthritic, neuroprotective,
XX antinflammatory, antiasthmatic, anti-allergic and cytostatic. This
XX polypeptide sequence is a single chain antibody that binds BlyS of the
XX invention. NOTE: The sequence data for this patent did not form part of
XX the printed specification, but was obtained in electronic format
XX directly from WIPO at ftp.wipo.int/pub/published pct_sequences.

Query Match 89.5%; Score 51; DB 7; Length 242;
Best Local Similarity 81.8%; Pred. No. 0.43;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
Db 156 QGDSLRITYAN 166

Search completed: April 28, 2005, 18:17:18
Job time : 89.8438 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2005, 18:06:50 ; Search time 22.3438 Seconds
(without alignments)
36.750 Million cell updates/sec

Title: US-10-088-639a-2_COPY_23_33
Perfect score: 57
Sequence: 1 QGDSLKTTYAS 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	54	94.7	104	3	US-08-793-450-2
2	54	94.7	238	3	US-08-793-450-6
3	50	87.7	97	2	US-08-665-202-35
4	50	87.7	97	4	US-09-315-574-35
5	50	87.7	101	4	US-09-726-219A-168
6	50	87.7	103	2	US-08-273-146-71
7	50	87.7	104	3	US-09-240-274-49
8	50	87.7	108	4	US-09-720-493-4
9	50	87.7	109	2	US-08-652-816A-16
10	50	87.7	109	2	US-08-665-202-34
11	50	87.7	109	4	US-09-315-574-34
12	50	87.7	111	4	US-09-203-768A-8
13	50	87.7	278	3	US-09-260-527-3
14	50	87.7	280	3	US-09-260-527-1
15	50	87.7	309	3	US-09-079-029-9
16	50	87.7	312	3	US-09-079-029-10
17	48	84.2	106	3	US-09-240-274-50
18	46	80.7	106	3	US-09-240-274-47
19	46	80.7	106	3	US-09-240-274-48
20	41	71.9	102	4	US-09-726-219A-174
21	40	70.2	125	4	US-09-471-276-1249
22	39	68.4	808	4	US-09-270-767-43054
23	38	66.7	117	4	US-09-530-139-20
24	38	66.7	224	4	US-09-530-139-38
25	38	66.7	238	4	US-09-530-139-36
26	38	66.7	624	4	US-09-538-092-227
27	36	63.2	185	4	US-09-562-737-117

28	35	61.4	86	4	US-09-134-000C-5678	Sequence 5678, Ap
29	35	61.4	200	4	US-09-252-991A-19243	Sequence 19243, A
30	34	59.6	118	4	US-09-513-999C-4175	Sequence 4175, Ap
31	34	59.6	119	2	US-08-652-816A-10	Sequence 10, Appl
32	34	59.6	139	4	US-09-471-276-837	Sequence 837, Appl
33	34	59.6	244	3	US-08-918-148-79	Sequence 79, Appl
34	34	59.6	244	4	US-09-138-091A-77	Sequence 77, Appl
35	34	59.6	376	4	US-09-328-352-4299	Sequence 4299, Ap
36	34	59.6	439	1	US-07-637-870-9	Sequence 9, Appli
37	34	59.6	439	1	US-07-637-399-6	Sequence 6, Appli
38	34	59.6	439	1	US-08-112-703-6	Sequence 6, Appli
39	34	59.6	657	3	US-08-893-852A-3	Sequence 3, Appli
40	34	59.6	657	3	US-08-821-818-3	Sequence 3, Appli
41	34	59.6	670	4	US-09-052-753B-3	Sequence 3, Appli
42	34	59.6	670	4	US-09-328-352-6725	Sequence 6725, Ap
43	34	59.6	812	4	US-09-538-092-643	Sequence 643, Ap
44	33	57.9	100	4	US-09-899-896-12	Sequence 12, Appl
45	33	57.9	149	4	US-09-270-767-40126	Sequence 40126, A

ALIGNMENTS

RESULT 1
US-08-793-450-2
; Sequence 2, Application US/08793450
; Patent No. 6312690
; GENERAL INFORMATION:
; APPLICANT: EDELMAN, LENA
; APPLICANT: MARGARITTE, CHRISTEL
; APPLICANT: KACZOREK, MICHEL
; APPLICANT: CHAABIRI, HASSAN
; TITLE OF INVENTION: MONOCLONAL RECOMBINANT ANTI-RHESUS D
; TITLE OF INVENTION:
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; ADDRESSER: P.C.
; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/793,450
; FILING DATE: 03-MAR-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 94/10566
; FILING DATE: 02-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 660-118-0 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 104 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-793-450-2

Query Match 94.7%; Score 54; DB 3; Length 104;
Best Local Similarity 90.9%; Pred. No. 0.0065;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
 |||||:||||
 Db 23 QGDSLRTYVAS 33

RESULT 2

US-08-793-450-6
 ; Sequence 6, Application US/08793450
 ; Patent No. 6312690
 ; GENERAL INFORMATION:
 ; APPLICANT: EDELMAN, LENA
 ; APPLICANT: MARGARITTE, CHRISTEL
 ; APPLICANT: KACZOREK, MICHEL
 ; APPLICANT: CHAABHI, HASSAN
 ; TITLE OF INVENTION: MONOCLONAL RECOMBINANT ANTI-RHESUS D
 ; TITLE OF INVENTION:
 ; NUMBER OF SEQUENCES: 25
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
 ; ADDRESSEE: P.C.
 ; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, SUITE 400
 ; CITY: ARLINGTON
 ; STATE: VA
 ; COUNTRY: USA
 ; ZIP: 22202
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/793,450
 ; FILING DATE: 03-MAR-1997
 ; CLASSIFICATION: 536
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: FR 94/10566
 ; FILING DATE: 02-SEP-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: OBLON, NORMAN F.
 ; REGISTRATION NUMBER: 24,618
 ; REFERENCE/DOCKET NUMBER: 660-118-0 PCT
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 703-413-3000
 ; TELEFAX: 703-413-2220
 ; INFORMATION FOR SEQ ID NO: 6:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 238 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-793-450-6

Query Match 94.7%; Score 54; DB 3; Length 238;
 Best Local Similarity 90.9%; Pred. No. 0.016;
 Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
 |||||:||||
 Db 42 QGDSLRTYVAS 52

RESULT 3

US-08-665-202-35
 ; Sequence 35, Application US/08665202
 ; Patent No. 5977322
 ; GENERAL INFORMATION:
 ; APPLICANT: Marks, James D.
 ; APPLICANT: Schier, Robert
 ; TITLE OF INVENTION: No. 5977322el High Affinity Human Antibodies to
 ; TITLE OF INVENTION: Tumor Antigens
 ; NUMBER OF SEQUENCES: 141
 ; CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP
 STREET: Two Embarcadero Center, Eighth Floor
 CITY: San Francisco
 STATE: California
 COUNTRY: USA
 ZIP: 94111-3834
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/665,202
 ; FILING DATE: 13-JUN-1996
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 60/000,238
 ; FILING DATE: 14-JUN-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 60/000,250
 ; FILING DATE: 15-JUN-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Hunter, Tom
 ; REGISTRATION NUMBER: 38,498
 ; REFERENCE/DOCKET NUMBER: 02307E-061410
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 576-0200
 ; TELEFAX: (415) 576-0300
 ; INFORMATION FOR SEQ ID NO: 35:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 97 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS:
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; US-08-665-202-35

Query Match 87.7%; Score 50; DB 2; Length 97;
 Best Local Similarity 81.8%; Pred. No. 0.035;
 Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
 |||||:||||
 Db 22 QGDSLRTYVAS 32

RESULT 4

US-09-315-574-35
 ; Sequence 35, Application US/09315574
 ; Patent No. 6512097
 ; GENERAL INFORMATION:
 ; APPLICANT: Marks, James D.
 ; APPLICANT: Schier, Robert
 ; TITLE OF INVENTION: No. 6512097el High Affinity Human Antibodies to
 ; TITLE OF INVENTION: Tumor Antigens
 ; NUMBER OF SEQUENCES: 141
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Majestic, Parsons, Siebert & Hsue P.C.
 ; STREET: Four Embarcadero Center, Suite 1100
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94111-4106
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/315,574
 ; FILING DATE: 20-MAY-99
 ; CLASSIFICATION: 530
 ; PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/000,238
FILING DATE: 14-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/000,250
FILING DATE: 15-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/665,202
FILING DATE: 13-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Hunter, Tom
REGISTRATION NUMBER: 38,498
REFERENCE/DOCKET NUMBER: 02307E-061411
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 97 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-315-574-35

Query Match 87.7%; Score 50; DB 4; Length 97;
Best Local Similarity 81.8%; Pred. No. 0.035;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
|||||:|||||
DB 22 QGDSLRSYYAS 32

RESULT 5

US-09-726-219A-168
Sequence 168, Application US/09726219A
Patent No. 6806079
GENERAL INFORMATION:
APPLICANT: Cambridge Antibody Technology
APPLICANT: Cambridge Antibody Technology Limited
APPLICANT: Medical Research Council
APPLICANT: McCafferty, John
APPLICANT: Pope, Anthony
APPLICANT: Johnson, Kevin
APPLICANT: Hoogenboom, Hendricus
APPLICANT: Grifiths, Andrew
APPLICANT: Jackson, Ronald
APPLICANT: Holliger, Kasper
APPLICANT: Marks, James
APPLICANT: Clackson, Timothy
APPLICANT: Chiswell, David
APPLICANT: Winter, Gregory
APPLICANT: Bonert, Timothy
TITLE OF INVENTION: Methods for Producing Members of Specific Binding Pairs
FILE REFERENCE: 213839-00013
CURRENT APPLICATION NUMBER: US/09/726,219A
CURRENT FILING DATE: 2000-11-28
PRIOR APPLICATION NUMBER: GB 9015198.6
PRIOR FILING DATE: 1990-07-10
PRIOR APPLICATION NUMBER: GB 9022845.3
PRIOR FILING DATE: 1990-10-19
PRIOR APPLICATION NUMBER: GB 9022845.3
PRIOR FILING DATE: 1990-10-19
PRIOR APPLICATION NUMBER: GB 9024503.6
PRIOR FILING DATE: 1990-11-12
PRIOR APPLICATION NUMBER: GB 9104744.9
PRIOR FILING DATE: 1991-03-06
PRIOR APPLICATION NUMBER: GB 9110549.4
PRIOR FILING DATE: 1991-05-15
PRIOR APPLICATION NUMBER: FCT/GB91/01134
PRIOR FILING DATE: 1991-07-10
PRIOR APPLICATION NUMBER: US 07/971,857
PRIOR FILING DATE: 1993-01-08

PRIOR APPLICATION NUMBER: US 08/484,893
FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 272
SOFTWARE: PatentIn version 3.1
SEQ ID NO 168
LENGTH: 101
TYPE: PRT
ORGANISM: Homo sapiens
US-09-726-219A-168

Query Match 87.7%; Score 50; DB 4; Length 101;
Best Local Similarity 81.8%; Pred. No. 0.037;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
|||||:|||||
DB 23 QGDSLRSYYAS 33

RESULT 6

US-08-273-146-71
Sequence 71, Application US/08273146
Patent No. 5855885
GENERAL INFORMATION:
APPLICANT: Smith, Rodger
APPLICANT: McCafferty, John
APPLICANT: Chiswell, David
APPLICANT: Darsley, Michael J.
APPLICANT: Fitzgerald, Kevin
APPLICANT: Kenten, John H.
APPLICANT: Martin, Mark T.
APPLICANT: Titmas, Richard C.
APPLICANT: Williams, Richard O.
TITLE OF INVENTION: The Isolation and Production of
Catalytic Antibodies using Phage Technology
NUMBER OF SEQUENCES: 71
TITLE OF INVENTION: Catalytic Antibodies using Phage Technology
CORRESPONDENCE ADDRESS:
ADDRESSEE: IGEN, Inc.
STREET: 1530 East Jefferson St.
CITY: Rockville
STATE: MD
COUNTRY: USA
ZIP: 20852
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/273,146
FILING DATE: 14-JUL-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ryan, John W.
REGISTRATION NUMBER: 33,771
REFERENCE/DOCKET NUMBER: 09000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-984-8000
TELEFAX: 301-230-0158
INFORMATION FOR SEQ ID NO: 71:
SEQUENCE CHARACTERISTICS:
LENGTH: 103 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-273-146-71

Query Match 87.7%; Score 50; DB 2; Length 103;
Best Local Similarity 81.8%; Pred. No. 0.037;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
|||||:|||||

```

Db      17 QGDSLSRYVAS 27

RESULT 7
US-09-240-274-49
; Sequence 49, Application US/09240274
; Patent No. 6255455
; GENERAL INFORMATION:
; APPLICANT: Siegel, Donald L.
; TITLE OF INVENTION: Rh(D)-BINDING PROTEINS AND MAGNETICALLY ACTIVATED CELL
; FILE REFERENCE: 09596-42U2
; CURRENT APPLICATION NUMBER: US/09/240,274
; CURRENT FILING DATE: 1999-01-29
; EARLIER APPLICATION NUMBER: 60/081,380
; EARLIER FILING DATE: 1998-04-10
; EARLIER APPLICATION NUMBER: 60/028,550
; EARLIER FILING DATE: 1996-10-11
; NUMBER OF SEQ ID NOS: 224
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 49
; TYPE: PRT
; LENGTH: 104
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: anti-Rh(D) chain J04
US-09-240-274-49

Query Match      87.7%; Score 50; DB 3; Length 104;
Best Local Similarity 81.8%; Pred. No. 0.038;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 QGDSLKYTVAS 11
        |||||:||||
Db      21 QGDSLSRYVAS 31

RESULT 8
US-09-720-493-4
; Sequence 4, Application US/09720493
; Patent No. 6827925
; GENERAL INFORMATION:
; APPLICANT: Cambridge Antibody Technology Limited
; APPLICANT: Williams, Andrew J
; APPLICANT: Tempest, Philip R
; APPLICANT: Holtet, Thor L
; APPLICANT: Main, Sarah H
; APPLICANT: Jackson, Helen
; APPLICANT: Daranola, Olalekan
; TITLE OF INVENTION: Improvements relating to antibodies
; FILE REFERENCE: AHB/CP5775333
; CURRENT APPLICATION NUMBER: US/09/720,493
; CURRENT FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: GB 9814383.7
; PRIOR FILING DATE: 1998-07-02
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-720-493-4

Query Match      87.7%; Score 50; DB 4; Length 108;
Best Local Similarity 81.8%; Pred. No. 0.039;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 QGDSLKYTVAS 11
        |||||:||||
Db      23 QGDSLSRYVAS 33

RESULT 9
US-08-652-816A-16
; Sequence 16, Application US/08652816A
; Patent No. 5872215
; GENERAL INFORMATION:
; APPLICANT: Osbourn, JK
; APPLICANT: Allen, DJ
; APPLICANT: McCafferty, JG
; TITLE OF INVENTION: Specific binding members, materials and
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/652,816A
; FILING DATE: 23-MAY-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9125579.4
; FILING DATE: 02-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9125579.8
; FILING DATE: 02-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9206318.9
; FILING DATE: 24-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9206372.6
; FILING DATE: 23-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9525004.9
; FILING DATE: 07-DEC-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9610824.6
; FILING DATE: 23-MAY-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB92/02240
; FILING DATE: 02-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/244,597
; FILING DATE: 01-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: David W. Clough
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 28111/33308
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 109 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-652-816A-16

Query Match      87.7%; Score 50; DB 2; Length 109;
Best Local Similarity 81.8%; Pred. No. 0.04;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 QGDSLKYTVAS 11
        |||||:||||
Db      23 QGDSLSRYVAS 33

RESULT 10
US-08-665-202-34
; Sequence 34, Application US/08665202

```

Patent No. 5977322
; GENERAL INFORMATION: James D.
; APPLICANT: Marks, Robert
; TITLE OF INVENTION: No. 5977322el High Affinity Human Antibodies to
; TITLE OF INVENTION: Tumor Antigens
; NUMBER OF SEQUENCES: 141
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/665,202
; FILING DATE: 13-JUN-1996
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/000,238
; FILING DATE: 14-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/000,250
; FILING DATE: 15-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Hunter, Tom
; REGISTRATION NUMBER: 38,498
; REFERENCE/DOCKET NUMBER: 02307E-061410
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 109 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-665-202-34

Query Match 87.7%; Score 50; DB 2; Length 109;
Best Local Similarity 81.8%; Pred. No. 0.04;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
|||:||||
Db 22 QGDSLRSYYAS 32

RESULT 11
US-09-315-574-34
; Sequence 34, Application US/09315574
; Patent No. 6512097
; GENERAL INFORMATION:
; APPLICANT: Marks, James D.
; APPLICANT: Schier, Robert
; TITLE OF INVENTION: No. 6512097el High Affinity Human Antibodies to
; TITLE OF INVENTION: Tumor Antigens
; NUMBER OF SEQUENCES: 141
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Majestic, Parsons, Siebert & Haue P.C.
; STREET: Four Embarcadero Center, Suite 1100
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4106
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/315,574
; FILING DATE: 20-MAY-99
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/000,238
; FILING DATE: 14-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/000,250
; FILING DATE: 15-JUN-1995
; APPLICATION NUMBER: US 08/665,202
; FILING DATE: 13-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hunter, Tom
; REGISTRATION NUMBER: 38,498
; REFERENCE/DOCKET NUMBER: 02307E-061411
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 109 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-315-574-34

Query Match 87.7%; Score 50; DB 4; Length 109;
Best Local Similarity 81.8%; Pred. No. 0.04;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
|||:||||
Db 22 QGDSLRSYYAS 32

RESULT 12
US-09-203-768A-8
; Sequence 8, Application US/09203768A
; Patent No. 6787638
; GENERAL INFORMATION:
; APPLICANT: Huse, William D.
; APPLICANT: Watkins, Jeffrey D.
; TITLE OF INVENTION: Tumor Specific Human Monoclonal Antibodies and Methods
; TITLE OF INVENTION: of Use
; FILE REFERENCE: P-IX 2947
; CURRENT APPLICATION NUMBER: US/09/203,768A
; CURRENT FILING DATE: 1998-12-02
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-203-768A-8

Query Match 87.7%; Score 50; DB 4; Length 111;
Best Local Similarity 81.8%; Pred. No. 0.04;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
|||:||||
Db 23 QGDSLRSYYAS 33

RESULT 13
US-09-260-527-3
; Sequence 3, Application US/09260527A

```
; Patent No. 6228599
; GENERAL INFORMATION:
; APPLICANT: Knox, J.P.
; APPLICANT: Mikkelsen, J.D.
; APPLICANT: Willats, W. G.
; TITLE OF INVENTION: ANTIBODY
; FILE REFERENCE: DY0U19.001AUS
; CURRENT APPLICATION NUMBER: US/09/260,527A
; CURRENT FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 278
; TYPE: PRT
; ORGANISM: UNKNOWN
; FEATURE:
; OTHER INFORMATION: Anti-homogalacturonan specific antibodies selected
; OTHER INFORMATION: from a naive phage display library known as the
; OTHER INFORMATION: Synthetic scFv Library (#1) from the Centre for
; OTHER INFORMATION: Protein Engineering, MRC Centre, Cambridge, UK
; US-09-260-527-3

Query Match      87.7%; Score 50; DB 3; Length 278;
Best Local Similarity 81.8%; Pred. No. 0.11;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 QGDSLKTYVAS 11
DB      175 QGDSLRSYYAS 185
|||||:|||||

RESULT 14
US-09-260-527-1
; Sequence 1, Application US/09260527A
; Patent No. 6228599
; GENERAL INFORMATION:
; APPLICANT: Knox, J.P.
; APPLICANT: Mikkelsen, J.D.
; APPLICANT: Willats, W. G.
; TITLE OF INVENTION: ANTIBODY
; FILE REFERENCE: DY0U19.001AUS
; CURRENT APPLICATION NUMBER: US/09/260,527A
; CURRENT FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 280
; TYPE: PRT
; ORGANISM: UNKNOWN
; FEATURE:
; OTHER INFORMATION: Anti-homogalacturonan specific antibodies from a
; OTHER INFORMATION: phage display library known as the Synthetic scFv
; OTHER INFORMATION: Library (#1) from the Centre for Protein
; OTHER INFORMATION: Engineering, MRC Centre, Cambridge, UK.
; US-09-260-527-1

Query Match      87.7%; Score 50; DB 3; Length 280;
Best Local Similarity 81.8%; Pred. No. 0.11;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 QGDSLKTYVAS 11
DB      177 QGDSLRSYYAS 187
|||||:|||||

RESULT 15
US-09-079-029-9
; Sequence 9, Application US/09079029
; Patent No. 6342369
; GENERAL INFORMATION:
; APPLICANT: Adams, Camilia W.
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Chuntharapai, Anan
```

```
; APPLICANT: Kim, Kyung J.
; TITLE OF INVENTION: Apo-2 Receptor
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/079,029
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Marschang, Diane L.
; REGISTRATION NUMBER: 35,600
; REFERENCE/DOCKET NUMBER: P1101R2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-5416
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 309 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; US-09-079-029-9

Query Match      87.7%; Score 50; DB 3; Length 309;
Best Local Similarity 81.8%; Pred. No. 0.12;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 QGDSLKTYVAS 11
DB      197 QGDSLRSYYAS 207
|||||:|||||

Search completed: April 28, 2005, 18:28:04
Job time : 23.3438 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2005, 18:24:46 ; Search time 68.2344 Seconds
(without alignments)

53.700 Million cell updates/sec

Title: US-10-088-639a-2_COPY_23_33

Perfect score: 57

Sequence: 1 QGDSLKTYVAS 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1426032 seqs, 333106140 residues

Total number of hits satisfying chosen parameters: 1426032

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
- 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
- 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
- 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	53	93.0	110	17	US-10-625-307A-18
2	53	93.0	253	10	US-09-880-748-1530
3	53	93.0	253	15	US-10-293-418-1530
4	51	89.5	242	10	US-09-880-748-2084
5	51	89.5	242	15	US-10-293-418-2084
6	51	89.5	243	10	US-09-880-748-2102
7	51	89.5	243	15	US-10-293-418-2102
8	51	89.5	247	10	US-09-880-748-2103
9	51	89.5	247	15	US-10-293-418-2103
10	51	89.5	248	10	US-09-880-748-2088
11	51	89.5	248	15	US-10-293-418-2088
12	51	89.5	249	10	US-09-880-748-2002
13	51	89.5	249	15	US-10-293-418-2002

14	51	89.5	250	10	US-09-880-748-1990	Sequence 1990, Ap
15	51	89.5	250	15	US-10-293-418-1990	Sequence 1990, Ap
16	50	87.7	96	15	US-10-308-817-106	Sequence 106, Ap
17	50	87.7	96	15	US-10-453-698-106	Sequence 106, Ap
18	50	87.7	101	16	US-10-803-622-168	Sequence 168, Ap
19	50	87.7	101	16	US-10-803-653-168	Sequence 168, Ap
20	50	87.7	104	10	US-09-848-798-49	Sequence 49, Appl
21	50	87.7	104	17	US-10-898-408-13	Sequence 13, Appl
22	50	87.7	108	17	US-10-727-155-94	Sequence 94, Appl
23	50	87.7	108	17	US-10-727-155-192	Sequence 192, Ap
24	50	87.7	108	17	US-10-727-155-316	Sequence 316, Ap
25	50	87.7	108	17	US-10-727-155-319	Sequence 319, Ap
26	50	87.7	108	17	US-10-890-945-4	Sequence 4, Appl
27	50	87.7	109	17	US-10-625-307A-41	Sequence 41, Appl
28	50	87.7	109	17	US-10-625-307A-65	Sequence 65, Appl
29	50	87.7	110	17	US-10-625-307A-43	Sequence 43, Appl
30	50	87.7	110	17	US-10-625-307A-63	Sequence 63, Appl
31	50	87.7	111	14	US-10-300-675-8	Sequence 8, Appl
32	50	87.7	111	14	US-10-300-675-16	Sequence 16, Appl
33	50	87.7	111	14	US-10-300-675-18	Sequence 18, Appl
34	50	87.7	111	14	US-10-300-675-20	Sequence 20, Appl
35	50	87.7	111	14	US-10-300-675-22	Sequence 22, Appl
36	50	87.7	111	14	US-10-300-675-24	Sequence 24, Appl
37	50	87.7	111	14	US-10-300-675-26	Sequence 26, Appl
38	50	87.7	111	15	US-10-032-037B-7	Sequence 7, Appl
39	50	87.7	111	15	US-10-029-988B-7	Sequence 7, Appl
40	50	87.7	111	15	US-10-032-423A-7	Sequence 7, Appl
41	50	87.7	111	15	US-10-029-926B-7	Sequence 7, Appl
42	50	87.7	111	16	US-10-910-124-8	Sequence 8, Appl
43	50	87.7	125	17	US-10-888-959-8	Sequence 8, Appl
44	50	87.7	153	10	US-09-988-115A-59	Sequence 59, Appl
45	50	87.7	229	14	US-10-310-719-36	Sequence 36, Appl

ALIGNMENTS

RESULT 1

US-10-625-307A-18

; Sequence 18, Application US/10625307A
; Publication No. US20050049403A1
; GENERAL INFORMATION:
; APPLICANT: Thompson, Julia E.
; APPLICANT: Vaughan, Tristan J.
; APPLICANT: Williams, Andrew J.
; APPLICANT: Green, Jonathan A.
; APPLICANT: Jackson, Ronald H.
; APPLICANT: Bacon, Louise
; APPLICANT: Johnson, Kevin S.
; APPLICANT: Wilton, Alison J.
; APPLICANT: Tempest, Philip R.
; APPLICANT: Pope, Anthony R.
; TITLE OF INVENTION: Specific Binding Members for Human Transforming Growth Factor Beta
; TITLE OF INVENTION: Materials and Methods
; FILE REFERENCE: 213839-00031
; CURRENT APPLICATION NUMBER: US/10/625,307A
; CURRENT FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: 09/054,847
; PRIOR FILING DATE: 1998-04-03
; PRIOR APPLICATION NUMBER: 08/571,755
; PRIOR FILING DATE: 1995-12-13
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 110
; TYPE: PRT
; ORGANISM: Human
US-10-625-307A-18

Query Match 93.0%; Score 53; DB 17; Length 110;
Best Local Similarity 90.9%; Pred. No. 0.054;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
QY      1 QGDSLKTYIAS 11
Db      23 QGDSLKSYAS 33

RESULT 2
US-09-880-748-1530
; Sequence 1530, Application US/09880748
; Publication No. US20030059937A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523
; CURRENT APPLICATION NUMBER: US/09/880,748
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 3239
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1530
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-748-1530

Query Match      93.0%; Score 53; DB 10; Length 253;
Best Local Similarity 90.9%; Pred. No. 0.13;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 QGDSLKTYIAS 11
Db      166 QGDSLKSYAS 176

RESULT 3
US-10-293-418-1530
; Sequence 1530, Application US/10293418
; Publication No. US20030223996A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523P2
; CURRENT APPLICATION NUMBER: US/10/293,418
; CURRENT FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: 60/331,469
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/340,817
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 09/880,748
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/212,210
; NUMBER OF SEQ ID NOS: 3247
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens

Query Match      93.0%; Score 53; DB 10; Length 253;
Best Local Similarity 90.9%; Pred. No. 0.13;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 QGDSLKTYIAS 11
Db      166 QGDSLKSYAS 176

RESULT 4
US-09-880-748-2084
; Sequence 2084, Application US/09880748
; Publication No. US20030059937A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523
; CURRENT APPLICATION NUMBER: US/09/880,748
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 3239
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2084
; LENGTH: 242
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-748-2084

Query Match      89.5%; Score 51; DB 10; Length 242;
Best Local Similarity 81.8%; Pred. No. 0.28;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 QGDSLKTYIAS 11
Db      156 QGDSLRTYAN 166

RESULT 5
US-10-293-418-2084
; Sequence 2084, Application US/10293418
; Publication No. US20030223996A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523P2
; CURRENT APPLICATION NUMBER: US/10/293,418
; CURRENT FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: 60/331,469
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/340,817
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 09/880,748
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/212,210
```

; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 2084
; LENGTH: 242
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-418-2084

Query Match 89.5%; Score 51; DB 15; Length 242;
Best Local Similarity 81.8%; Pred. No. 0.28;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
Db 156 QGDSLRTYYAN 166
|||||:||||:

RESULT 6

US-09-880-748-2102
; Sequence 2102, Application US/09880748
; Publication No. US20030059937A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.

; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys

; FILE REFERENCE: PF523
; CURRENT APPLICATION NUMBER: US/09/880,748

; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,210

; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/240,816

; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276,248

; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/277,379

; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293,499

; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 3239

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2102

; LENGTH: 243
; TYPE: PRT

; ORGANISM: Homo sapiens
US-09-880-748-2102

Query Match 89.5%; Score 51; DB 10; Length 243;
Best Local Similarity 81.8%; Pred. No. 0.28;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
Db 157 QGDSLRTYYAN 167
|||||:||||:

RESULT 7

US-10-293-418-2102
; Sequence 2102, Application US/10293418
; Publication No. US20030223996A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.

; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys

; FILE REFERENCE: PF523P2
; CURRENT APPLICATION NUMBER: US/10/293,418

; PRIOR FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: 60/331,469

; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/340,817

; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 09/880,748

; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/293,499

; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/277,379

; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 2102

; LENGTH: 243
; TYPE: PRT

; ORGANISM: Homo sapiens
US-10-293-418-2102

Query Match 89.5%; Score 51; DB 15; Length 243;
Best Local Similarity 81.8%; Pred. No. 0.28;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
Db 157 QGDSLRTYYAN 167
|||||:||||:

RESULT 8

US-09-880-748-2103

; Sequence 2103, Application US/09880748
; Publication No. US20030059937A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.

; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys

; FILE REFERENCE: PF523
; CURRENT APPLICATION NUMBER: US/09/880,748

; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,210

; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/240,816

; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276,248

; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/277,379

; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293,499

; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 3239

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2103

; LENGTH: 247
; TYPE: PRT

; ORGANISM: Homo sapiens
US-09-880-748-2103

Query Match 89.5%; Score 51; DB 10; Length 247;
Best Local Similarity 81.8%; Pred. No. 0.29;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
Db 161 QGDSLRTYYAN 171
|||||:||||:

RESULT 9

US-10-293-418-2103

; Sequence 2103, Application US/10293418
; Publication No. US20030223996A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.

; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys

; FILE REFERENCE: PF523P2
; CURRENT APPLICATION NUMBER: US/10/293,418

; PRIOR FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: 60/331,469

; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/340,817

```

; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 09/880,748
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 2103
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-418-2103

```

```

Query Match      89.5%; Score 51; DB 15; Length 247;
Best Local Similarity 81.8%; Pred. No. 0.29;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 QGDSLKTYVAS 11
      |||||:||||:
Db      161 QGDSLRTYYAN 171

```

RESULT 10

```

US-09-880-748-2088
; Sequence 2088, Application US/09880748
; Publication No. US20030059937A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; CURRENT APPLICATION NUMBER: US/09/880,748
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 3239
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2088
; LENGTH: 248
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-748-2088

```

```

Query Match      89.5%; Score 51; DB 10; Length 248;
Best Local Similarity 81.8%; Pred. No. 0.29;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 QGDSLKTYVAS 11
      |||||:||||:
Db      162 QGDSLRTYYAN 172

```

RESULT 11

```

US-10-293-418-2088
; Sequence 2088, Application US/10293418
; Publication No. US20030223996A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys

```

```

; FILE REFERENCE: PF523P2
; CURRENT APPLICATION NUMBER: US/10/293,418
; CURRENT FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: 60/331,469
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/340,817
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 09/880,748
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 2088
; LENGTH: 248
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-418-2088

```

```

Query Match      89.5%; Score 51; DB 15; Length 248;
Best Local Similarity 81.8%; Pred. No. 0.29;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 QGDSLKTYVAS 11
      |||||:||||:
Db      162 QGDSLRTYYAN 172

```

RESULT 12

```

US-09-880-748-2002
; Sequence 2002, Application US/09880748
; Publication No. US20030059937A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523
; CURRENT APPLICATION NUMBER: US/09/880,748
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 3239
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2002
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-748-2002

```

```

Query Match      89.5%; Score 51; DB 10; Length 249;
Best Local Similarity 81.8%; Pred. No. 0.29;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 QGDSLKTYVAS 11
      |||||:||||:
Db      163 QGDSLRTYYAN 173

```

RESULT 13

US-10-293-418-2002
; Sequence 2002, Application US/10293418
; Publication No. US20030223996A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523P2
; CURRENT APPLICATION NUMBER: US/10/293,418
; CURRENT FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: 60/331,469
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/340,817
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 09/880,748
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 2002
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-418-2002

Query Match 89.5%; Score 51; DB 15; Length 249;
Best Local Similarity 81.8%; Pred. No. 0.29;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QGDSLKTYAS 11
Db 163 QGDSLRTYYAN 173

RESULT 14
US-09-880-748-1990
; Sequence 1990, Application US/09880748
; Publication No. US20030059937A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523
; CURRENT APPLICATION NUMBER: US/09/880,748
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 3239
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1990
; LENGTH: 250
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-748-1990

Query Match 89.5%; Score 51; DB 10; Length 250;
Best Local Similarity 81.8%; Pred. No. 0.29;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QGDSLKTYAS 11
Db 164 QGDSLRTYYAN 174

RESULT 15
US-10-293-418-1990
; Sequence 1990, Application US/10293418
; Publication No. US20030223996A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523P2
; CURRENT APPLICATION NUMBER: US/10/293,418
; CURRENT FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: 60/331,469
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/340,817
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 09/880,748
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1990
; LENGTH: 250
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-418-1990

Query Match 89.5%; Score 51; DB 15; Length 250;
Best Local Similarity 81.8%; Pred. No. 0.29;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QGDSLKTYAS 11
Db 164 QGDSLRTYYAN 174

Search completed: April 28, 2005, 19:02:12
Job time : 69.2344 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2005, 18:05:55 ; Search time 15.125 Seconds
(without alignments)
69.976 Million cell updates/sec

Title: US-10-088-639A-2_COPY_23_33
Perfect score: 57
Sequence: 1 QGDSLKTYVAS 11
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	53	93.0	109	2 S38496	Ig lambda chain -
2	50	87.7	96	2 S36060	Ig lambda chain -
3	50	87.7	106	2 S38495	Ig lambda chain -
4	50	87.7	108	2 S38498	Ig lambda chain -
5	50	87.7	108	2 S47184	Ig lambda chain -
6	50	87.7	109	2 S19663	Ig lambda chain V
7	50	87.7	110	2 S36272	Ig lambda chain V
8	50	87.7	115	2 S13726	Ig lambda chain V
9	50	87.7	127	2 S70444	Ig lambda chain pr
10	49	86.0	233	2 S25748	Ig lambda chain -
11	45	78.9	233	2 S25741	Ig lambda chain -
12	44	77.2	2135	2 T14602	variant-specific s
13	41	71.9	110	2 S19672	Ig lambda chain V
14	39	68.4	1729	2 T18396	erythrocyte membra
15	39	68.4	2324	2 T18378	variant-specific s
16	38	66.7	429	2 E64617	hypothetical prote
17	38	66.7	429	2 F71896	hypothetical prote
18	38	66.7	624	2 S44938	nitrogen permease
19	37	64.9	146	2 S02083	Ig lambda chain V-
20	36	63.2	270	2 AG9916	probable hydrolase
21	36	63.2	452	1 WHHUF	phenylalanine 4-no
22	36	63.2	547	2 D84942	methionine-tRNA 11
23	36	63.2	1159	1 A44280	inner layer protel
24	35	61.4	197	2 B83497	hypothetical prote
25	35	61.4	269	2 H84378	hypothetical prote
26	35	61.4	313	2 H72724	probable microsoma
27	35	61.4	367	2 T24058	hypothetical prote
28	35	61.4	408	2 AC2188	hypothetical prote
29	35	61.4	561	2 JE0158	dwarf protein, OSD

30	35	61.4	580	2 S72211	N-acetyl-beta-D-gl
31	35	61.4	868	2 AF3204	autotransporter pr
32	35	61.4	2228	2 T14029	variant-specific s
33	34	59.6	108	1 L3HUSH	Ig lambda chain V-
34	34	59.6	161	2 B4462	allophycocyanin be
35	34	59.6	189	2 D71978	hypothetical prote
36	34	59.6	189	2 E64527	hypothetical prote
37	34	59.6	218	2 G70438	hypothetical prote
38	34	59.6	284	2 T09452	vdIC protein - Hel
39	34	59.6	439	2 E86813	xylose isomerase (
40	34	59.6	446	2 T03267	probable phosphop
41	34	59.6	446	2 T02221	phosphopyruvate hy
42	34	59.6	448	1 WMBP82	gene 8 protein - p
43	34	59.6	464	2 F97807	sodium/pantothanat
44	34	59.6	529	2 A84260	hypothetical prote
45	34	59.6	561	2 T36084	hypothetical prote

ALIGNMENTS

RESULT 1

S38496
Ig lambda chain - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 21-Jan-2000
C:Accession: S38496
R:Marks, J.D.; Owehand, W.H.; Bye, J.M.; Finnern, R.; Gorick, B.D.; Voak, D.; Thorpe, S.
Submitted to the EMBL Data Library, June 1993
A:Description: Human antibody fragments specific for human blood group antigens from a p
A:Reference number: S38488
A:Accession: S38496
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-109 <MAR>
A:Cross-references: EMBL:223031; NID:g414039; PIDN:CAA80566.1; PID:g414040
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-89/Domain: immunoglobulin homology <IMM>

Query Match 93.0%; Score 53; DB 2; Length 109;
Best Local Similarity 90.9%; Pred. No. 0.004;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QGDSLKTYVAS 11
|||||:||||
Db 23 QGDSLKSYAS 33

RESULT 2

S36060
Ig lambda chain - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 22-Nov-1993 #sequence_revision 01-Dec-1995 #text_change 21-Jan-2000
C:Accession: S36060
R:Williams, S.C.
Submitted to the EMBL Data Library, April 1993
A:Reference number: S36046
A:Accession: S36060
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-96 <WIL>
A:Cross-references: EMBL:222202; NID:g312325; PIDN:CAA80211.1; PID:g312326
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-89/Domain: immunoglobulin homology <IMM>

Query Match 87.7%; Score 50; DB 2; Length 96;
Best Local Similarity 81.8%; Pred. No. 0.014;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QGDSLKTYVAS 11
|||||:||||

Db 23 QGDSLRSYYAS 33

RESULT 3

S38495
IG lambda chain - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 21-Jan-2000
C:Accession: S38495
R:Marks, J.D.; Ouweland, W.H.; Bye, J.M.; Finnern, R.; Gorick, B.D.; Voak, D.; Thorpe, S.
submitted to the EMBL Data Library, June 1993
A:Description: Human antibody fragments specific for human blood group antigens from a patient
A:Reference number: S38488
A:Accession: S38495
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-106 <MAR>
A:Cross-references: EMBL:223029; NID:g414037; PIDN:CAA80564.1; PID:g414038
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-89/Domain: immunoglobulin homology <IMM>

Query Match 87.7%; Score 50; DB 2; Length 106;
Best Local Similarity 81.8%; Pred. No. 0.015;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
|||||:||||
Db 23 QGDSLRSYYAS 33

RESULT 4

S38498
IG lambda chain - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 21-Jan-2000
C:Accession: S38498
R:Marks, J.D.; Ouweland, W.H.; Bye, J.M.; Finnern, R.; Gorick, B.D.; Voak, D.; Thorpe, S.
submitted to the EMBL Data Library, June 1993
A:Description: Human antibody fragments specific for human blood group antigens from a patient
A:Reference number: S38488
A:Accession: S38498
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-108 <MAR>
A:Cross-references: EMBL:223035; NID:g414043; PIDN:CAA80570.1; PID:g414044
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:14-88/Domain: immunoglobulin homology <IMM>

Query Match 87.7%; Score 50; DB 2; Length 108;
Best Local Similarity 81.8%; Pred. No. 0.016;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
|||||:||||
Db 22 QGDSLRSYYAS 32

RESULT 5

S47184
IG lambda chain - human
C:Species: Homo sapiens (man)
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 21-Jan-2000
C:Accession: S47184
R:McIntosh, R.S.; Tandon, N.; Metcalfe, R.A.; Weetman, A.P.
submitted to the EMBL Data Library, June 1994
A:Description: Cloning and analysis of IgM anti-thyroglobulin autoantibodies from patient
A:Reference number: S47181
A:Accession: S47184
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-108 <MC1>

A:Cross-references: EMBL:X79783; NID:g506426; PIDN:CAA56179.1; PID:g506427
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-89/Domain: immunoglobulin homology <IMM>

Query Match 87.7%; Score 50; DB 2; Length 108;
Best Local Similarity 81.8%; Pred. No. 0.016;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
|||||:||||
Db 23 QGDSLRSYYAS 33

RESULT 6

S19663
IG lambda chain V region (clone alpha-BSA3) - human
C:Species: Homo sapiens (man)
C>Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 20-Jun-2000
C:Accession: S19663
R:Marks, J.D.; Hoogenboom, H.R.; Bonnert, T.P.; McCafferty, J.; Griffiths, A.D.; Winter, J.
Mol. Biol. 222, 581-597, 1991
A:Title: By-passing immunization. Human antibodies from V-gene libraries displayed on phage
A:Reference number: S19663; MUID:92085276; PMID:1748994
A:Accession: S19663
A:Molecule type: mRNA
A:Residues: 1-109 <MAR>
A:Cross-references: EMBL:X61640; NID:g29492; PIDN:CAA43821.1; PID:g1340166
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-89/Domain: immunoglobulin homology <IMM>

Query Match 87.7%; Score 50; DB 2; Length 109;
Best Local Similarity 81.8%; Pred. No. 0.016;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
|||||:||||
Db 23 QGDSLRSYYAS 33

RESULT 7

S36272
IG lambda chain V region (clone alpha-THY-29) - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 21-Jan-2000
C:Accession: S36272
R:Griffiths, A.D.; Malmqvist, M.; Marks, J.D.; Bye, J.M.; Embleton, M.J.; McCafferty, J.;
EMBO J. 12, 725-734, 1993
A:Title: Human anti-self antibodies with high specificity from phage display libraries.
A:Reference number: S36256; MUID:93178448; PMID:7679990
A:Accession: S36272
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-110 <GR1>
A:Cross-references: EMBL:Z18833; NID:g33419; PIDN:CAA79285.1; PID:g939912
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-89/Domain: immunoglobulin homology <IMM>

Query Match 87.7%; Score 50; DB 2; Length 110;
Best Local Similarity 81.8%; Pred. No. 0.016;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
|||||:||||
Db 23 QGDSLRSYYAS 33

RESULT 8

S13726
IG lambda chain V region - human
C:Species: Homo sapiens (man)

C>Date: 25-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 21-Jan-2000
C:Accession: S13726
R:Frippiat, J.P.; Chuchana, P.; Bernhard, F.; Buluwela, L.; Lefranc, G.; Lefranc, M.P.
Nucleic Acids Res. 18, 7134, 1990
A:Title: First genomic sequence of a human Ig variable lambda gene belonging to subgroup A
A:Reference number: S13726; MUID:91088295; PMID:2124677
A:Accession: S13726
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-115 <PRI>
A:Cross-references: EMBL:X56178; NID:g33404; PIDN:CAA39639.1; PID:g33405
C:Genetics: 16/1
A:introns: 16/1
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:34-108/Domain: immunoglobulin homology <IMM>

Query Match 87.7%; Score 50; DB 2; Length 115;
Best Local Similarity 81.8%; Pred. No. 0.017;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
Db 42 QGDSLRSYYAS 52

RESULT 9
S70444
Ig lambda chain precursor V region - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004
C:Accession: S70444; S70426
R:Cuisinier, A.M.; Fumoux, F.; Fougereau, M.; Tonnelle, C.
Mol. Immunol. 29, 1363-1373, 1992
A:Title: IGM kappa/lambda EBV human B cell clone: an early step of differentiation of B
A:Reference number: S70444; MUID:93024508; PMID:1383695
A:Accession: S70444
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-127 <CUI>
A:Cross-references: UNIPROT:QNSD6
A:Experimental source: clone E29.1
R:Tonnelle, C.
submitted to the EMBL Data Library, May 1990
A:Reference number: S70426
A:Accession: S70426
A:Molecule type: mRNA
A:Residues: 1-90 <TON>
A:Cross-references: EMBL:X53070
A:Experimental source: cell line E29.1, clone VL 29-1
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-127/Product: Ig lambda chain V region (fragment) #status predicted <MAT>
F:34-108/Domain: immunoglobulin homology <IMM>

Query Match 87.7%; Score 50; DB 2; Length 127;
Best Local Similarity 81.8%; Pred. No. 0.018;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
Db 42 QGDSLRSYYAS 52

RESULT 10
S25748
Ig lambda chain - human
C:Species: Homo sapiens (man)
C>Date: 22-Nov-1993 #sequence_revision 26-May-1995 #text_change 21-Jan-2000
C:Accession: S25748
R:Combrinato, G.; Klobeck, H.G.
Eur. J. Immunol. 21, 1513-1522, 1991

A:Title: V(lambda) and J(lambda)-C(lambda) gene segments of the human immunoglobulin lam
A:Reference number: S16439; MUID:91257162; PMID:1904362
A:Accession: S25748
A:Status: preliminary; translation not shown
A:Molecule type: mRNA
A:Residues: 1-233 <COM>
A:Cross-references: EMBL:X57813; NID:g33725; PIDN:CAA40950.1; PID:g33726
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:148-216/Domain: immunoglobulin homology <IMM>

Query Match 86.0%; Score 49; DB 2; Length 233;
Best Local Similarity 81.8%; Pred. No. 0.055;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
Db 42 QGDSLRSYYAS 52

RESULT 11
S25741
Ig lambda chain - human
C:Species: Homo sapiens (man)
C>Date: 22-Nov-1993 #sequence_revision 26-May-1995 #text_change 21-Jan-2000
C:Accession: S25741
R:Combrinato, G.; Klobeck, H.G.
Eur. J. Immunol. 21, 1513-1522, 1991
A:Title: V(lambda) and J(lambda)-C(lambda) gene segments of the human immunoglobulin lam
A:Reference number: S16439; MUID:91257162; PMID:1904362
A:Accession: S25741
A:Status: preliminary; translation not shown
A:Molecule type: mRNA
A:Residues: 1-233 <COM>
A:Cross-references: EMBL:X57805; NID:g33707; PIDN:CAA40943.1; PID:g33708
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:148-216/Domain: immunoglobulin homology <IMM>

Query Match 78.9%; Score 45; DB 2; Length 233;
Best Local Similarity 63.6%; Pred. No. 0.34;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
Db 42 QGDTIRSYAS 52

RESULT 12
T14602
variant-specific surface protein - malaria parasite (Plasmodium falciparum) (fragments)
C:Species: Plasmodium falciparum
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C:Accession: T14602
R:Voss, T.S.; Felger, I.; Weiss, N.; Beck, H.P.
submitted to the EMBL Data Library, February 1998
A:Description: Identification of a conserved 5' flanking region of Plasmodium falciparum
A:Reference number: Z18158
A:Accession: T14602
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2135 <VOS>
A:Cross-references: UNIPROT:O61077; EMBL:AF050740; NID:g2944094; PID:g2944095; PIDN:AA001
C:Genetics:
A:Gene: varph17

Query Match 77.2%; Score 44; DB 2; Length 2135;
Best Local Similarity 77.8%; Pred. No. 5.4;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTY 9
Db 160 EGDSTKTY 168

RESULT 13

S19672
IG lambda chain V region (clone alpha-TEL14) - human
C:Species: Homo sapiens (man)
C>Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 20-Jun-2000
C:Accession: S19672
R:Marka, J.D.; Hoogenboom, H.R.; Bonnert, T.P.; McCafferty, J.; Griffiths, A.D.; Winter, J. Mol. Biol. 222, 581-597, 1991
A:Title: By-passing immunization. Human antibodies from V-gene libraries displayed on phage
A:Reference number: S19663; MUID:92085276; PMID:1748994
A:Accession: S19672
A:Molecule type: mRNA
A:Residues: 1-110 <NAR>
A:Cross-references: EMBL:X61644; NID:G37856; PIDN:CAA43825.1; PID:G1335384
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-89/Domain: immunoglobulin homology <IMM>

Query Match 71.9%; Score 41; DB 2; Length 110;
Best Local Similarity 72.7%; Pred. NO. 0.96;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 QGDSLKTYYS 11
DB 23 QGDSLSRSYAS 33
:||||:|
:||||:|

RESULT 14

T18396
erythrocyte membrane protein variant 2 - malaria parasite (Plasmodium falciparum) (fragment)
C:Species: Plasmodium falciparum
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T18396
R:Baruch, D.I.; Pasloske, B.L.; Singh, H.B.; Bi, X.; Ma, X.C.; Feldman, M.; Taraschi, T. Cell 82, 77-87, 1995
A:Title: Cloning the P. falciparum gene encoding PfEMP1, a malarial variant antigen and
A:Reference number: Z18925; MUID:95330812; PMID:7541722
A:Accession: T18396
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-1729 <BAR>
A:Cross-references: UNIPROT:Q25734; EMBL:U27339; NID:G914920; PID:G914921; PIDN:AAA89134
C:Genetics:
A:Gene: EMP1
A>Note: var-2

Query Match 68.4%; Score 39; DB 2; Length 1729;
Best Local Similarity 66.7%; Pred. NO. 43;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYTY 9
DB 151 EGDSEIKTHY 159
:||||:|
:||||:|

RESULT 15

T18378
variant-specific surface protein 1 - malaria parasite (Plasmodium falciparum)
N:Alternate names: erythrocyte membrane binding protein 1 (EMP1)
C:Species: Plasmodium falciparum
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T18378
R:Baruch, D.I.; Pasloske, B.L.; Singh, H.B.; Bi, X.; Ma, X.C.; Feldman, M.; Taraschi, T. Cell 82, 77-87, 1995
A:Title: Cloning the P. falciparum gene encoding PfEMP1, a malarial variant antigen and
A:Reference number: Z18925; MUID:95330812; PMID:7541722
A:Accession: T18378
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-2924 <BAR>
A:Cross-references: UNIPROT:Q25733; EMBL:U27338; NID:G914918; PID:G914919; PIDN:ARB60251

C:Genetics:

A:Gene: EMP1

A:Introns: 2476/3

Query Match 68.4%; Score 39; DB 2; Length 2924;
Best Local Similarity 66.7%; Pred. NO. 74;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYTY 9

:||||:|

DB 151 EGDSEIKTHY 159

Search completed: April 28, 2005, 18:25:49
Job time : 16.125 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 01:42:46 ; Search time 9.13208 Seconds
(without alignments)
62.721 Million cell updates/sec

Title: US-10-088-639A-2_COPY_23_33
Perfect score: 57
Sequence: 1 QGDSLKTYVAS 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	38	66.7	624	1 NPR2 YEAST	P39923 saccharomyc
2	36	63.2	270	1 DLHH_SALTI	Q82368 salmonella
3	36	63.2	270	1 DLHH_SALTY	Q916m9 salmonella
4	36	63.2	306	1 B3GI DROME	O97422 d galactosy
5	36	63.2	452	1 PH4H HUMAN	P00439 homo sapien
6	36	63.2	545	1 PH4H HUMAN	O93hd7 buchnera ap
7	36	63.2	547	1 SYM_EUCAL	P57210 buchnera ap
8	36	63.2	1159	1 RRPO_ROTGI	P35942 rotavirus (
9	34	59.6	108	1 LV3A HUMAN	P01714 homo sapien
10	34	59.6	161	1 PHAB_SYNY3	Q01952 synchocyst
11	34	59.6	439	1 XYL4_LACLA	O9c5g7 lactococcus
12	34	59.6	446	1 ENO2_MAIZE	P42895 zea mays (m
13	34	59.6	446	1 ENO_ORYZA	Q42971 oryza sativ
14	34	59.6	448	1 HEAD_BPPZA	P07531 bacterioph
15	34	59.6	586	1 LREI YEAST	P17579 saccharomyc
16	34	59.6	657	1 MY16 MOUSE	P17564 mus musculu
17	34	59.6	812	1 IF38 YEAST	P32497 saccharomyc
18	33	57.9	145	1 HBBI_TRICR	P10785 triturus cr
19	33	57.9	161	1 PHAB_SYNY4	Q02524 synchocyst
20	33	57.9	178	1 YQAC_BACSU	P45900 bacillus su
21	33	57.9	268	1 KSGA_PSEAE	Q915u5 p dimethyla
22	33	57.9	270	1 RECX_LACLA	O9c5g7 lactococcus
23	33	57.9	289	1 TTHR_CHICK	P25324 gallus gall
24	33	57.9	307	1 PRTR_TRIAL	P23653 tritirachiu
25	33	57.9	446	1 ENO1_MAIZE	P26301 zea mays (m
26	33	57.9	448	1 HEAD_BPPH2	P13849 bacterioph
27	33	57.9	451	1 MB31_ARATH	O04309 arabidopsi
28	33	57.9	453	1 PH4H_MOUSE	P16331 mus musculu
29	33	57.9	453	1 PH4H_RAT	P04176 rattus norv
30	33	57.9	558	1 MNT2 YEAST	P53059 saccharomyc
31	33	57.9	670	1 AT6A_HUMAN	P18850 homo sapien
32	33	57.9	736	1 TREE_MAGGR	O42622 magnaporthe
33	33	57.9	751	1 GCP2_CABEL	P91406 caenorhabdi

Q10850	mycobacteri	1327	1	YK06_MYCTU	57.9	33	34
Q04952	saccharomyc	1785	1	GLS3 YEAST	57.9	33	35
Q8y231	ralstonia s	77	1	FISL RALSO	56.1	32	36
P17616	bacillus su	114	1	YPUD_BACSU	56.1	32	37
Q30881	synchococc	117	1	CYCG_SYNP2	56.1	32	38
Q9xt48	macropus eu	201	1	TNFB_MACEU	56.1	32	39
P28676	homo sapien	217	1	GRAN_HUMAN	56.1	32	40
Q98qt6	mycoplasma	224	1	Y295_MYCPU	56.1	32	41
Q01074	bacterioph	229	1	VG07_BPP22	56.1	32	42
P03019	escherichia	250	1	FNR_ECOLI	56.1	32	43
Q9aq50	klebsiella	250	1	FNR_KLBOX	56.1	32	44
Q82787	salmonella	250	1	FNR_SALTI	56.1	32	45
P37428	salmonella	250	1	FNR_SALTY	56.1	32	46
Q91a24	shigella dy	250	1	FNR_SHIDY	56.1	32	47
O8ae82	yersinia pe	250	1	FNR_YERPE	56.1	32	48
O8x814	escherichia	271	1	DLHH_ECO57	56.1	32	49
P56262	escherichia	271	1	DLHH_ECOLI	56.1	32	50
Q98487	rhizobium l	279	1	KSGA_RHILO	56.1	32	51
Q82d92	yersinia pe	343	1	CPDP_YERPE	56.1	32	52
Q92j43	rickettsia	395	1	ODO2_RICCN	56.1	32	53
P47323	mycoplasma	407	1	OPPB_MYCGE	56.1	32	54
O82845	tetragenoco	435	1	XYLA_TETHA	56.1	32	55
P50743	bacillus su	436	1	ENGA_BACSU	56.1	32	56
Q37888	bacterioph	449	1	HEAD_BBP03	56.1	32	57
P26745	bacterioph	475	1	VG2_BPLP7	56.1	32	58
O26157	methanobact	499	1	VG02_BPP22	56.1	32	59
Q39085	arabidopsi	553	1	SYE_METHH	56.1	32	60
P93472	pisum sativ	561	1	DIM_ARATH	56.1	32	61
P42396	bacillus su	567	1	DIM_PEA	56.1	32	62
P15624	saccharomyc	569	1	YXID_BACSU	56.1	32	63
P44092	haemophilus	594	1	SYFB YEAST	56.1	32	64
P10105	drosophila	622	1	PPID_HABIN	56.1	32	65
O8xwb1	ralstonia s	635	1	HMLA_DROME	56.1	32	66
P38339	saccharomyc	636	1	MUTL_RALSO	56.1	32	67
Q39214	arabidopsi	666	1	RGD1 YEAST	56.1	32	68
P11939	gallus gall	926	1	RPWL_ARATH	56.1	32	69
O51134	borrelia bu	367	1	FOS_CHICK	55.3	31.5	70
Q9pex5	xylella fas	145	1	NUSB BORBU	54.4	31	71
O87ap2	xylella fas	161	1	YH81_XYLFT	54.4	31	72
O01513	caenorhabdi	161	1	YN31_XANCP	54.4	31	73
Q37875	bacterioph	183	1	YS4L_CABEL	54.4	31	74
P77288	escherichia	185	1	LYCV_BPP1	54.4	31	75
Q9yeK4	aeropyrum p	187	1	YFCV_ECOLI	54.4	31	76
O8d7y6	streptococc	225	1	Y574_AERPE	54.4	31	77
P16562	homo sapien	242	1	YK70_STRMU	54.4	31	78
Q8K9K6	streptococc	243	1	CRS2_HUMAN	54.4	31	79
Q99xq1	streptococc	243	1	YH90_STRP3	54.4	31	80
Q8n237	streptococc	243	1	YL63_STRP8	54.4	31	81
Q60477	cavia porce	243	1	CRS2_CAVPO	54.4	31	82
P821w3	streptococc	251	1	TPIS_STRP8	54.4	31	83
P82478	streptococc	251	1	TPIS_STRPY	54.4	31	84
P21911	schizosacch	252	1	DH8B_SCHPO	54.4	31	85
Q8e644	streptococc	252	1	TPIS_STRAS	54.4	31	86
Q8e644	streptococc	252	1	TPIS_STRAS	54.4	31	87
P72484	streptococc	252	1	TPIS_STRMU	54.4	31	88
Q97pn1	streptococc	252	1	TPIS_STRPN	54.4	31	89
Q8vvc1	streptococc	252	1	TPIS_STRTR	54.4	31	90
P45199	haemophilus	257	1	FNR_HABIN	54.4	31	91
P46233	vibrio para	293	1	MOTY_VIBPA	54.4	31	92
Q7tlv56	prochloroc	300	1	PRMA_PROMP	54.4	31	93
Q9hlyv3	thermoplasm	338	1	PUR5_THEAC	54.4	31	94
Q9x276	thermotoga	350	1	REDD_STRCO	54.4	31	95
Q9K1K6	vibrio chol	357	1	BUKI_THEMA	54.4	31	96
Q62376	mus musculu	364	1	LUXP_VIBCH	54.4	31	97
Q9fd15	zymomonas m	367	1	HIS8_WOLSU	54.4	31	98
Q97ft6	clostridium	378	1	RUI7_MOUSE	54.4	31	99
Q8cny4	staphylococ	416	1	PUR2_ZYMMO	54.4	31	100
P08621	homo sapien	431	1	TIG_CLOAB	54.4	31	101
Q9req3	zymomonas m	433	1	TIG_STAEP	54.4	31	102
		437	1	RUI7_HUMAN	54.4	31	103
		448	1	RMUC_ZYMMO	54.4	31	104
							105
							106

```

107 31 54.4 448 1 YG4Q YEAST
108 31 54.4 471 1 RUI7 XENLA
109 31 54.4 478 1 VTNC MOUSE
110 31 54.4 487 1 ENGA CHLPN
111 31 54.4 488 1 TY3H ANGAN
112 31 54.4 509 1 SRC1 BOVIN
113 31 54.4 509 1 SRC1 FIG
114 31 54.4 518 1 BAE2 HUMAN
115 31 54.4 574 1 PHF1 CLOPA
116 31 54.4 621 1 Y708 CHLMU
117 31 54.4 694 1 PKK CAMJE
118 31 54.4 744 1 HXCI HAEIN
119 31 54.4 748 1 Y875 METJA
120 31 54.4 758 1 METE SNEEL
121 31 54.4 790 1 YDDB ECOLI
122 31 54.4 817 1 DLG3 HUMAN
123 31 54.4 833 1 CAPA YERPE
124 31 54.4 849 1 DLG3 MOUSE
125 31 54.4 849 1 DLG3 RAT
126 31 54.4 883 1 DPOL WHV8
127 31 54.4 884 1 DPOL WHV7
128 31 54.4 884 1 PDB1 ARATH
129 31 54.4 967 1 PDB1 ARATH
130 31 54.4 980 1 POLG LIV
131 31 54.4 1200 1 DP3A CAMJE
132 31 54.4 1241 1 TRK1 SACBA
133 31 54.4 1249 1 APAF MOUSE
134 31 54.4 1520 1 ACDF VIBCH
135 31 54.4 1597 1 PTBE HUMAN
136 31 54.4 3412 1 POLG TREVS
137 31 54.4 3414 1 POLG TREVH
138 31 54.4 3414 1 POLG TREVW
139 31 54.4 3564 1 CSM1 MOUSE
140 31 54.4 3565 1 CSM1 HUMAN
141 31 54.4 4910 1 MDN1 YEAST
142 30.5 53.5 424 1 COA2 BPFD
143 30.5 53.5 424 1 COA2 BPM13
144 30 52.6 115 1 RL19 THETN
145 30 52.6 120 1 RET3 CHICK
146 30 52.6 136 1 RET3 HUMAN
147 30 52.6 136 1 RET3 MOUSE
148 30 52.6 145 1 HBB2 TRICR
149 30 52.6 153 1 YMEF ECOLI
150 30 52.6 161 1 Y064_XANAC

```

ALIGNMENTS

```

RESULT 1
NPR2 YEAST
ID NPR2 YEAST STANDARD; PRT; 624 AA.

```

```

AC P3923;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Nitrogen permease regulator 2.
OS NPR2 OR YEL062W.

```

```

OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c;
RX MEDLINE=95172238; PubMed=7867803;

```

```

RA Rousset G., Simon M., Ripoch P., Buhler J.-M.;
RT "A second nitrogen permease regulator in Saccharomyces cerevisiae.";

```

```

RL FEBS Lett. 359:215-219 (1995).
RN [2]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RC STRAIN=97313264; PubMed=9169868;
RX

```

```

RA Dietrich F.S., Mulligan J.T., Hennessy K.M., Yelton M.A., Allen E.,
RA Araujo R., Aviles E., Berno A., Brennan T., Carpenter J., Chen E.,
RA Cherry J.M., Chung E., Duncan M., Guzman E., Hartzell G.,
RA Hunnicke-Smith S., Hyman R.W., Kayser A., Komp C., Lashkari D., Lew H.,
RA Lin D., Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P.,
RA Oh C., Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T.,
RA Smith V., Taylor P., Wei Y., Botstein D., Davis R.W.;
RA "The nucleotide sequence of Saccharomyces cerevisiae chromosome V.";
RT Nature 387:78-81 (1997).
RL
CC -!- FUNCTION: Post-transcriptional regulator of nitrogen permeases. It
CC could be involved in the putative NPR1-dependent phosphorylation
CC of nitrogen permeases or in the processing and targeting of the
CC nitrogen permeases at the level of the endoplasmic reticulum.
CC -!- INDUCTION: By urea and proline.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X79105; CAA55721.1; -
CC EMBL; U18795; AAB55025.1; -
CC F01R; S44938; S44938.
CC GenBank; U18795; S44938.
CC TRANSFAC; T04834; -
CC SGD; S000788; NPR2.
CC GO; GO:0015824; P-L-proline transport; IMP.
CC GO; GO:0015840; Piurea transport; IMP.
CC GO; GO:0015840; Piurea transport; IMP.
CC Transcription regulation.
CC DOMAIN 363 429 SER-RICH.
CC
CC SEQUENCE 624 AA; 70888 MW; 6349863463C0FC5D CRC64;

```

```

DR EMBL; X79105; CAA55721.1; -
DR EMBL; U18795; AAB55025.1; -
DR F01R; S44938; S44938.
DR GenBank; U18795; S44938.
DR TRANSFAC; T04834; -
DR SGD; S000788; NPR2.
DR GO; GO:0015824; P-L-proline transport; IMP.
DR GO; GO:0015840; Piurea transport; IMP.
DR GO; GO:0015840; Piurea transport; IMP.
DR Transcription regulation.
DR DOMAIN 363 429 SER-RICH.
DR
DR SEQUENCE 624 AA; 70888 MW; 6349863463C0FC5D CRC64;

```

```

Query Match 66.7%; Score 38; DB 1; Length 624;
Best Local Similarity 63.6%; Pred. No. 9;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 QGDSLKTYVAS 11
|||:|||||
Db 456 QGOTLKTYVES 466

```

RESULT 2

```

DLHH SALTII
ID DLHH SALTII STANDARD; PRT; 270 AA.

```

```

AC Q823B8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Putative carboxymethylenebutenolidase (EC 3.1.1.45) (Dienelactone
DE hydrolase) (DLH).
DE YSGA OR STV3592 OR T3330.

```

```

GN Salmonella typhi.
OS Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;

```

```

RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebatina M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogan A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M.A., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrall B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
enterica serovar Typhi CT18.";

```

```

RL Nature 413:848-852 (2001).

```

[2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=Ty2 / ATCC 700931;
 RX MEDLINE=22531367; PubMed=12644504;
 RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
 RA Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.,
 RT "Comparative genomics of *Salmonella enterica* serovar Typhi strains Ty2
 RT and CT18.";
 RL J. Bacteriol. 185:2330-2337(2003).
 CC -!- CATALYTIC ACTIVITY: 4-carboxymethylenebut-2-en-4-olide + H(2)O = 4
 CC oxohex-2-enedioate.
 CC -!- SIMILARITY: Belongs to the dienelactone hydrolase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AL627278; CAD07925.1; -;
 DR EMBL; AE016845; AAO70858.1; -;
 DR InterPro; IPR002925; DLH.
 DR Pfam; PF01738; DLH; 1.
 KW Hypothetical protein; Hydrolase; Complete proteome.
 FT ACT_SITE 147 147 BY SIMILARITY.
 FT ACT_SITE 204 204 BY SIMILARITY.
 FT ACT_SITE 236 236 BY SIMILARITY.
 FT CONFLICT 201 201 S -> G (IN REF. 2).
 SQ SEQUENCE 270 AA; 29092 MW; 0D7630B6046075B2 CRC64;
 Query Match 63.2%; Score 36; DB 1; Length 270;
 Best Local Similarity 60.0%; Pred. No. 9.3;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 1 QGDSLKTYVA 10
 DB ||| : |||
 38 QGDDMPAYYA 47
 RESULT 3
 ID DLH SALTY STANDARD; PRT; 270 AA.
 AC Q9L6W9;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DE Putative carboxymethylenebutenolidase (EC 3.1.1.45) (Dienelactone
 DE hydrolase) (DLH).
 GN YSGA OR STM3967 OR STM1.22.
 OS *Salmonella typhimurium*.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; *Salmonella*.
 OX NCBI_TaxID=602;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
 RX MEDLINE=21534948; PubMed=11677609;
 RA McClelland M., Sanderson K.B., Spieth J., Clifton S.W., Latreille P.,
 RA Courtney L., Porwollik S., Ali J., Dante M., Du P., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 RT "Complete genome sequence of *Salmonella enterica* serovar Typhimurium
 RT LT2.";
 RL Nature 413:852-856(2001).
 CC -!- CATALYTIC ACTIVITY: 4-carboxymethylenebut-2-en-4-olide + H(2)O = 4
 CC oxohex-2-enedioate.
 CC -!- SIMILARITY: Belongs to the dienelactone hydrolase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF233324; AAF33425.1; -;
 DR EMBL; AE008885; AAL22811.1; -;
 DR StyGene; SG?????; YSGA.
 DR InterPro; IPR002925; DLH.
 DR Pfam; PF01738; DLH; 1.
 KW Hypothetical protein; Hydrolase; Complete proteome.
 FT ACT_SITE 147 147 BY SIMILARITY.
 FT ACT_SITE 204 204 BY SIMILARITY.
 FT ACT_SITE 236 236 BY SIMILARITY.
 SQ SEQUENCE 270 AA; 29062 MW; 1C2374B6047DA872 CRC64;
 Query Match 63.2%; Score 36; DB 1; Length 270;
 Best Local Similarity 60.0%; Pred. No. 9.3;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 1 QGDSLKTYVA 10
 DB ||| : |||
 38 QGDDMPAYYA 47
 RESULT 4
 ID B3GI DROME STANDARD; PRT; 306 AA.
 AC Q97422; Q8IRS6; Q95SR5;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 15-MAR-2004 (Rel. 43, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Galactosylgalactosylxylosylprotein 3-beta-glucuronosyltransferase I
 DE (EC 2.4.1.135) (Beta-1,3-glucuronyltransferase I) (UDPGlcUA:Gal Beta-1,3-Gal-R
 DE glucuronyltransferase-I) (DmgGlcAT-I).
 DE Glucuronyltransferase (GLCAT-I).
 GN GLCAT-I OR BCDNA:GH05057 OR EG:EG0007.5 OR CG32775.
 OS *Drosophila melanogaster* (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; *Drosophila*.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A., FUNCTION, COPACTOR, AND DEVELOPMENTAL STAGE.
 RX MEDLINE=22513909; PubMed=12511570;
 RA Kim B.-T., Tsuchida K., Lincecum J., Kitagawa K., Bernfield M.,
 RA Sugahara K.;
 RT "Identification and characterization of three *Drosophila melanogaster*
 RT glucuronyltransferases responsible for the synthesis of the conserved
 RT glycosaminoglycan-protein linkage region of proteoglycans: two novel
 RT homologs exhibit broad specificity toward oligosaccharides from
 RT proteoglycans, glycoproteins, and glycosphingolipids.";
 RL J. Biol. Chem. 278:9116-9124(2003).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Oregon-R;
 RX MEDLINE=20196011; PubMed=10731137;
 RA Benos P.V., Gatt M.K., Ashburner M., Murphy L., Harris D.,
 RA Barrell B.G., Ferraz C., Vidal S., Brun C., Demailles J., Cadieu E.,
 RA Dreano S., Gloux S., Lelaure V., Mottier S., Galibert F., Borikova D.,
 RA Minana B., Kafatos F.C., Louis C., Siden-Kiamos I., Bolshakov S.,
 RA Papagiannakis G., Spanos L., Cox S., Madueno E., de Pablo B.,
 RA Modellell J., Peter A., Schoettler P., Werner M., Mourikioti F.,
 RA Beinert N., Dowe G., Schaefer U., Jaekle H., Bucheton A.,
 RA Callister D.M., Campbell L.A., Darlamitsou A., Henderson N.S.,
 RA McMillan P.J., Salles C., Tait E.A., Valenti P., Saunders R.D.C.,
 RA Glover D.M.;
 RT "From sequence to chromosome: the tip of the X chromosome of *D.*
 RT *melanogaster*.";
 RL Science 287:2220-2222(2000).
 RN [3]
 RP SEQUENCE FROM N.A.

CC STRAIN=Berkeley;
 CC MEDLINE=20196006; PubMed=10731132;
 CC RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 CC RA Amatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,
 CC RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 CC RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 CC RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
 CC RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.B.G.,
 CC RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 CC RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 CC RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 CC RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 CC RA Burtis K.C., Busam D.A., Butler H., Cadiou E., Center A., Chandra I.,
 CC RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 CC RA de Fabios B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 CC RA Dodson K.B., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 CC RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 CC RA Fostler C., Gabrielian A.E., Gang N.S., Gelbart W.M., Glasser K.,
 CC RA Glodek A., Gong F., Correll J.H., Gu Z., Guan P., Harris M.,
 CC RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
 CC RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 CC RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 CC RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 CC RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 CC RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 CC RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 CC RA Mount S.M., Moy M., Murphy S., Murphy L., Muzny D.M., Nelson D.L.,
 CC RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleib J.M.,
 CC RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 CC RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 CC RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 CC RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 CC RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 CC RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 CC RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 CC RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 CC RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 CC RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 CC RT "The genome sequence of *Drosophila melanogaster*.";
 CC RL Science 287:2185-2195(2000).
 CC [4]
 CC RP REVISIONS.
 CC RX MEDLINE=22426069; PubMed=12537572;
 CC RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 CC RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
 CC RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 CC RA Bettencourt B.R., Celinker S.E., de Grey A.D.N.J., Drysdale R.A.,
 CC RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.O.,
 CC RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 CC RA Lewis S.E.;
 CC RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a
 CC RT systematic review.";
 CC RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
 CC [5]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=Berkeley; TISSUE=Head;
 CC RX MEDLINE=22426066; PubMed=12537569;
 CC RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,
 CC RA George R.A., Guarin H., Kronmiller B., Pacleib J.M., Park S., Wan K.H.,
 CC RA Rubin G.M., Celinker S.E.;
 CC RT "A *Drosophila* full-length cDNA resource.";
 CC RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).
 CC CC -!- FUNCTION: Involved in the biosynthesis of Lz/HNK-1 carbohydrate
 CC CC epitope on both glycolipids and glycoproteins. Shows strict
 CC CC specificity for Gal beta1-3Gal beta1-4Xyl, exhibiting negligible
 CC CC incorporation into other galactoside substrates.
 CC CC -!- CATALYTIC ACTIVITY: UDP-glucuronate + 3-beta-D-galactosyl-4-
 CC CC beta-D-galactosyl-O-beta-D-xylosylprotein = UDP + 3-beta-D-
 CC CC glucuronosyl-3-beta-D-galactosyl-4-beta-D-galactosyl-O-beta-D-
 CC CC xylosylprotein.
 CC CC -!- COFACTOR: Manganese.
 CC CC -!- PATHWAY: Glycosylation.
 CC CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Golgi (Potential).
 CC

CC -!- DEVELOPMENTAL STAGE: Expressed at low levels from early embryos to
 CC adults; maximal expression in third instar larvae.
 CC -!- SIMILARITY: Belongs to the glycosyltransferase family 43.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AB080695; BAC55095.1; -
 CC EMBL; AL033125; CA21824.1; ALT_SEQ.
 CC EMBL; AF003430; AAN09117.1; -
 CC EMBL; AY060634; AAL28182.1; -
 CC HSSP; O94766; 1FGG.
 CC FlyBase; FBgn0066114; GlcAT-I.
 CC InterPro; IPR005027; Glyco trans 43.
 CC Pfam; PF03360; Glyco_transf_43; 1.
 CC Transferrase; Glycoprotein; Transmembrane; Signal-anchor; Golgi stack;
 CC Manganese; Multigene family.
 CC DOMAIN 1 11 CYTOPLASMIC (POTENTIAL).
 CC TRANSMEM 12 29 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
 CC (POTENTIAL).
 CC DOMAIN 30 306 LUMENAL (POTENTIAL).
 CC ACT SITE 252 252 CATALYTIC BASE (BY SIMILARITY).
 CC METAL 163 163 MANGANESE (BY SIMILARITY).
 CC CARBOHYD 90 90 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CONFLICT 55 55 V -> L (IN REF. 1 AND 2).
 CC SEQUENCE 306 AA; 35071 MW; 69910A46534218B0 CRC64;
 CC
 CC Query Match 63.2%; Score 36; DB 1; Length 306;
 CC Best Local Similarity 70.0%; Pred. No. 11;
 CC Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 CC
 CC QY 1 QGDSLKTYYA 10
 CC DB |||||
 CC 45 QGDTLPTIYA 54
 CC
 CC RESULT 5
 CC PH4H HUMAN STANDARD; PRT; 452 AA.
 CC ID P04339; Q16717; Q8TC14;
 CC AC 21-JUL-1986 (Rel. 01, Created)
 CC DT 21-JUL-1986 (Rel. 01, Last sequence update)
 CC DT 10-OCT-2003 (Rel. 42, Last annotation update)
 CC DE Phenylalanine-4-hydroxylase (EC 1.14.16.1) (PAH) (Phe-4-
 CC DE monooxygenase).
 CC GN PAH.
 CC OS Homo sapiens (Human).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 CC OX NCBI_TaxID=9606;
 CC RN [1] SEQUENCE FROM N.A.
 CC RP TISSUE=Liver;
 CC RX MEDLINE=85199778; PubMed=2986678;
 CC RA Kwok S.C.M., Ledley F.D., Dilella A.G., Robson K.J.H., Woo S.L.C.;
 CC RT "Nucleotide sequence of a full-length complementary DNA clone and
 CC RT amino acid sequence of human phenylalanine hydroxylase.";
 CC RL Biochemistry 24:556-561(1985).
 CC RN [2]
 CC RP SEQUENCE FROM N.A.
 CC RA Scriver C.R., Nowacki P.M., Byck S., Prevost L.;
 CC RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 CC RN [3]
 CC RP SEQUENCE FROM N.A.
 CC RC TISSUE=Liver;
 CC RX MEDLINE=22388257; PubMed=12477932;
 CC RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 CC RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zebberg B., Buetow K.H., Schaefer C.P., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedlin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinaki M.I., Skalska U., Smalins D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 117-424.
RX MEDLINE=98069646; PubMed=9406548;
RA Erlandsen H., Fusetti F., Martinez A., Hough E., Flatmark T.,
RA Stevens R.C.;
RT "Crystal structure of the catalytic domain of human phenylalanine
RT hydroxylase reveals the structural basis for phenylketonuria";
RL Nat. Struct. Biol. 4:995-1000(1997).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 117-424.
RX MEDLINE=99060040; PubMed=9843368;
RA Erlandsen H., Flatmark T., Stevens R.C., Hough E.;
RT "Crystallographic analysis of the human phenylalanine hydroxylase
RT catalytic domain with bound catechol inhibitors at 2.0-A resolution";
RL Biochemistry 37:15638-15646(1998).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS) OF 117-452.
RX MEDLINE=98307935; PubMed=9642259;
RA Fusetti F., Erlandsen H., Flatmark T., Stevens R.C.;
RT "Structure of tetrameric human phenylalanine hydroxylase and its
RT implications for phenylketonuria";
RL J. Biol. Chem. 273:16962-16967(1998).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 118-424.
RX MEDLINE=20160497; PubMed=10694386;
RA Erlandsen H., Bjorbo E., Flatmark T., Stevens R.C.;
RT "Crystal structure and site-specific mutagenesis of pterin-bound human
RT phenylalanine hydroxylase";
RL Biochemistry 39:2208-2217(2000).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS) OF 103-427.
RX MEDLINE=21576340; PubMed=11718561;
RA Andersen O.A., Flatmark T., Hough E.;
RT "High resolution crystal structures of the catalytic domain of human
RT phenylalanine hydroxylase in its catalytically active Fe(II) form and
RT binary complex with tetrahydrobiopterin";
RL J. Mol. Biol. 314:266-278(2001).
RN [9]
RP REVIEW ON PKU VARIANTS.
RX MEDLINE=91348681; PubMed=1679029;
RA Konecki D.S., Lichter-Konecki U.;
RT "The phenylketonuria locus: current knowledge about alleles and
RT mutations of the phenylalanine hydroxylase gene in various
RT populations";
RL Hum. Genet. 87:377-388(1991).
RN [10]
RP REVIEW ON PKU VARIANTS.
RX MEDLINE=91061429; PubMed=2246858;
RA Cotton R.G.;
RT "Heterogeneity of phenylketonuria at the clinical, protein and DNA
RT levels";
RL J. Inher. Metab. Dis. 13:739-750(1990).
RN [11]
RP REVIEW ON PKU VARIANTS.
RX MEDLINE=93244826; PubMed=1301187;

RA Eisenmith R.C., Woo S.L.C.;
RT "Molecular basis of phenylketonuria and related
RT hyperphenylalaninemia: mutations and polymorphisms in the human
RT phenylalanine hydroxylase gene";
RL Hum. Mutat. 1:13-22(1992).
RN [12]
RP DATABASE OF PKU VARIANTS.
RX MEDLINE=96174613; PubMed=8594560;
RA Hoang L., Byck S., Prevost L., Scriber C.R.;
RT "PAH Mutation Analysis Consortium Database: a database for disease-
RT producing and other allelic variation at the human PAH locus";
RL Nucleic Acids Res. 24:127-131(1996).
RN [13]
RP VARIANTS PKU PRO-311.
RX MEDLINE=88294030; PubMed=2840952;
RA Lichter-Konecki U., Konecki D.S., Dilella A.G., Brayton K., Marvit J.,
RA Hahn T.M., Trefz F.K., Woo S.L.C.;
RT "Phenylalanine hydroxylase deficiency caused by a single base
RT substitution in an exon of the human phenylalanine hydroxylase
RT gene";
RL Biochemistry 27:2881-2885(1988).
RN [14]
RP VARIANTS PKU LYS-280.
RX MEDLINE=89190664; PubMed=2564729;
RA Lyonnet S., Caillaud C., Rey F., Berthelon M., Frezal J., Rey J.,
RA Munnich A.;
RT "Molecular genetics of phenylketonuria in Mediterranean countries: a
RT mutation associated with partial phenylalanine hydroxylase
RT deficiency";
RL Am. J. Hum. Genet. 44:511-517(1989).
RN [15]
RP VARIANTS PKU PRO-311.
RX MEDLINE=90136055; PubMed=2615649;
RA Hofman K.J., Antonarakis S.E., Missiou-Tsangarakis S., Boehm C.D.,
RA Valle D.;
RT "Phenylketonuria in the Greek population. Haplotype analysis of the
RT phenylalanine hydroxylase gene and identification of a PKU
RT mutation";
RL Mol. Biol. Med. 6:245-250(1989).
RN [16]
RP VARIANTS PKU LEU-364 DEL.
RX MEDLINE=90368081; PubMed=1975559;
RA Svensson E., Andersson B., Hagenfeldt L.;
RT "Two mutations within the coding sequence of the phenylalanine
RT hydroxylase gene";
RL Hum. Genet. 85:300-304(1990).
RN [17]
RP VARIANTS PKU GLN-261.
RX MEDLINE=91150775; PubMed=1671810;
RA Dianzani I., Forrest S.M., Camaschella C., Saglio G., Ponzzone A.,
RA Cotton R.G.;
RT "Screening for mutations in the phenylalanine hydroxylase gene from
RT Italian patients with phenylketonuria by using the chemical cleavage
RT method: a new splice mutation";
RL Am. J. Hum. Genet. 48:631-635(1991).
RN [18]
RP VARIANTS PKU SER-255.
RX MEDLINE=91196738; PubMed=2014802;
RA Hofman K.J., Steel G., Kazanian H.H., Valle D.;
RT "Phenylketonuria in U.S. blacks: molecular analysis of the
RT phenylalanine hydroxylase gene";
RL Am. J. Hum. Genet. 48:791-798(1991).
RN [19]
RP VARIANTS PKU TRP-252 AND LEU-281.
RX MEDLINE=91169521; PubMed=1672294;
RA Okano Y., Wang T., Eisenmith R.C., Longhi R., Riva E., Giovannini M.,
RA Cerone R., Romano C., Woo S.L.C.;
RT "Phenylketonuria missense mutations in the Mediterranean";
RL Genomics 9:96-103(1991).
RN [20]
RP VARIANTS PKU LEU-281.
RX MEDLINE=91169506; PubMed=1672290;
RA Dworniczak B., Grudka K., Stumper J., Bartholome K.,

RA Aulehla-Scholz C., Horst J.;
 RT "Phenylalanine hydroxylase gene: novel missense mutation in exon 7
 RL causing severe phenylketonuria.";
 RL Genomics 9:193-199(1991).
 RN [21]
 RN VARIANTS PKU SER-48 AND GLY-221.
 RX MEDLINE=911348682; PubMed=1679030;
 RA Konecki D.S., Schlottner M., Trefz F.K., Lichter-Konecki U.;
 RT "The identification of two mis-sense mutations at the PAH gene locus
 RT in a Turkish patient with phenylketonuria.";
 RL Hum. Genet. 87:389-393(1991).
 RN [22]
 RN VARIANT PKU ILE-94 DEL.
 RX MEDLINE=91236693; PubMed=1709636;
 RA Caillaud C., Lyonnet S., Rey F., Melle D., Frebourg T., Berthelon M.,
 RA Vilarinho L., Vaz Osorio R., Rey J., Munnich A.;
 RT "A 3-base pair in-frame deletion of the phenylalanine hydroxylase
 RT gene results in a kinetic variant of phenylketonuria.";
 RL J. Biol. Chem. 266:9351-9354(1991).
 RN [23]
 RN VARIANTS NON-PKU HPA VAL-306 AND ASN-415.
 RX MEDLINE=93052278; PubMed=1358789;
 RA Economou-Petersen E., Henriksen K.F., Guldberg P., Guettler F.;
 RT "Molecular basis for nonphenylketonuria hyperphenylalaninemia.";
 RT "Molecular basis for nonphenylketonuria hyperphenylalaninemia.";
 Query Match 63.2%; Score 36; DB 1; Length 452;
 Best Local Similarity 70.0%; Pred. No. 16;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 QGDSLKTYVA 10
 DB 336 QGDSIKAYCA 345
 |||||
 RESULT 6
 ID SYM_BUCAP STANDARD; PRT; 545 AA.
 AC Q9ZHD7;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Methionyl-tRNA synthetase (EC 6.1.1.10) (Methionine--tRNA ligase)
 DE (MetRS)
 GN METG OR BUSG102.
 OS Buchnera aphidicola (subsp. Schizaphis graminum).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Buchnera.
 OX NCBI_TaxID=98794;
 RN [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=22084549; PubMed=12089438;
 RA Tamas I., Klaason L., Canbaeck B., Naeslund A.K., Eriksson A.-S.,
 RA Werngreen J.J., Sandstrom J.P., Moran N.A., Andersson S.G.E.;
 RT "50 million years of genomic stasis in endosymbiotic bacteria.";
 RL Science 296:2376-2379(2002).
 RN [2]
 RN SEQUENCE OF 1-234 FROM N.A.
 RX MEDLINE=98440331; PubMed=9767718;
 RA Clark M.A., Baumann L., Baumann P.;
 RA "Buchnera aphidicola (Aphid endosymbiont) contains genes encoding
 RT enzymes of histidine biosynthesis.";
 RL Curr. Microbiol. 37:356-358(1998).
 CC -!- FUNCTION: Is required not only for elongation of protein synthesis
 CC but also for the initiation of all mRNA translation through
 CC initiator tRNA(fMet) aminoacylation (By similarity).
 CC -!- CATALYTIC ACTIVITY: ATP + L-methionine + tRNA(Met) = AMP +
 CC diphosphate + L-methionyl-tRNA(Met).
 CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
 CC -!- SUBUNIT: Monomer (By similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.
 CC MetG subfamily 1.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AE014085; AM67672.1; -;
 DR EMBL; AF067228; AAC97364.1; -;
 DR HSSP; P00959; 1MEA.
 DR HAMAP; MF 00098; -; 1.
 DR InterPro; IPR002300; tRNA-synt_1a.
 DR InterPro; IPR001412; tRNA-synt_1.
 DR InterPro; IPR002304; tRNA-synt_met.
 DR Pfam; PF00133; tRNA-synt_1; 1.
 DR PRINTS; PR01041; TRNASYNTHMET.
 DR TIGRFAMs; TIGR00398; metG; 1.
 DR PROSITE; PS00178; AA_TRNA_LIGASE_I; 1.
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 KW Metal-binding; Zinc; Complete proteome.
 FT SITE 15 25
 FT SITE 332 336
 FT BINDING 335 335
 FT BINDING 335 335
 FT METAL 146 146
 FT METAL 149 149
 FT METAL 159 159
 FT METAL 162 162
 FT METAL 162 162
 SQ SEQUENCE 545 AA; 63822 MW; 77F79377EF21492A CRC64;
 Query Match 63.2%; Score 36; DB 1; Length 545;
 Best Local Similarity 77.8%; Pred. No. 20;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 3 DSLKTYVAS 11
 DB 353 DSLRYYYAS 361
 |||||
 RESULT 7
 ID SYM_BUCAI STANDARD; PRT; 547 AA.
 AC P57210;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Methionyl-tRNA synthetase (EC 6.1.1.10) (Methionine--tRNA ligase)
 DE (MetRS).
 GN METG OR BUI109.
 OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
 OS symbiotic bacterium).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Buchnera.
 OX NCBI_TaxID=118099;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=Tokyo 1998;
 RX MEDLINE=20445173; PubMed=10993077;
 RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
 RT "Genome sequence of the endocellular bacterial symbiont of aphids
 RT Buchnera sp. APS.";
 RL Nature 407:81-86(2000).
 CC -!- FUNCTION: Is required not only for elongation of protein synthesis
 CC but also for the initiation of all mRNA translation through
 CC initiator tRNA(fMet) aminoacylation (By similarity).
 CC -!- CATALYTIC ACTIVITY: ATP + L-methionine + tRNA(Met) = AMP +
 CC diphosphate + L-methionyl-tRNA(Met).
 CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
 CC -!- SUBUNIT: Monomer (By similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.
 CC MetG subfamily 1.
 CC -----

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
CC or send an email to license@isb-sib.ch.
CC -----

DR EMBL; AP001118; BAB12828.1; --
DR HAMAP; P00959; IMEA.
DR InterPro; MF_00098; --; 1.
DR InterPro; IPR002300; tRNA-synt_1a.
DR InterPro; IPR001412; tRNA-synt_1.
DR InterPro; IPR002304; tRNA-synt_met.
DR Pfam; PF00133; tRNA-synt_1; 1.
DR PRINTS; PR01041; TRNASYNTHMET.
DR TIGRFAMs; TIGR00398; metG; 1.
DR PROSITE; PS00178; AA_TRNA_LIGASE_I; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Metal-binding; Zinc; Complete proteome.
FT SITE 15 25 "HIGH" REGION.
FT SITE 332 336 "KMSKS" REGION.
FT BINDING 335 335 ATP (BY SIMILARITY).
FT METAL 146 146 ZINC (BY SIMILARITY).
FT METAL 149 149 ZINC (BY SIMILARITY).
FT METAL 159 159 ZINC (BY SIMILARITY).
FT METAL 162 162 ZINC (BY SIMILARITY).
SQ SEQUENCE 547 AA; 64282 MW; PFE802670A15EB7 CRC64;

Query Match 63.2%; Score 36; DB 1; Length 547;
Best Local Similarity 77.8%; Pred. No. 20;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 3 DSLKTYAS 11
DB 353 DSLRYTAS 361
|||||

RESULT 8
RRPO ROTGI STANDARD; PRT; 1159 AA.
AC P35942;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE RNA-directed RNA polymerase subunit VP1 (EC 2.7.7.48) (Inner layer
DE protein VP1) (Core protein VP1).
GN S1.
OS Rotavirus (group B / strain IDIR) (IDIR agent).
OC Viruses; dsRNA viruses; Reoviridae; Rotavirus.
OX NCBI_TaxID=28877;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=93297100; PubMed=8390749;
RA Eiden J.J., Hirshon C.;
RT "Sequence analysis of group B rotavirus gene 1 and definition of a
RT rotavirus-specific sequence motif within the RNA polymerase gene.";
RL Virology 152:154-160(1993).
CC -!- FUNCTION: RNA POLYMERASE ACTIVITY IS THOUGHT TO BE ASSOCIATED
CC WITH VP1.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA} (N).
CC -!- SUBCELLULAR LOCATION: Inner layer.
CC -!- SIMILARITY: Belongs to the rotaviruses VP1 protein family.
CC -----

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
CC or send an email to license@isb-sib.ch.
CC -----

DR EMBL; M97203; --; NOT ANNOTATED_CDS.
DR PIR; A44280; A44280.
DR InterPro; IPR007095; RNA pol DS PS.
DR InterPro; IPR007097; RNA_pol_tel.
DR InterPro; IPR008817; Rotavirus VP1.
DR Pfam; PF05740; Rotavirus_VP1; 1.
KW Core protein; RNA-directed RNA polymerase; Transferase.
SQ SEQUENCE 1159 AA; 131649 MW; E355F9BF79E225A8 CRC64;

Query Match 63.2%; Score 36; DB 1; Length 1159;
Best Local Similarity 70.0%; Pred. No. 45;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 GDSLKTYAS 11
DB 1039 GNTLXTYLAS 1048
|:|||||

RESULT 9
LV3A HUMAN STANDARD; PRT; 108 AA.
ID LV3A HUMAN
AC P01714;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ig lambda chain V-III region SH.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE.
RX MEDLINE=70166723; PubMed=4909564;
RA Titani K., Wikler M., Shinoda T., Putnam F.W.;
RT "The amino acid sequence of a lambda type Bence-Jones protein. 3. The
RT complete amino acid sequence and the location of the disulfide
RT bridges.";
RL J. Biol. Chem. 245:2171-2176(1970).
CC -!- MISCELLANEOUS: This is a Bence-Jones protein.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like domain.
DR PIR; A01980; L3HUSH.
DR HSSP; P80748; 2LOI.
DR GO; GO:0005576; C:extracellular; NAS.
DR GO; GO:0003823; P:antigen binding; NAS.
DR GO; GO:0006955; P:immune response; NAS.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGv; 1.
DR PROSITE; PS50835; IG LIKE; 1.
KW Immunoglobulin V region; Bence-Jones protein.
FT DOMAIN 1 97 IG-LIKE.
FT DISULFID 21 86
FT NON_TER 108 108
SQ SEQUENCE 108 AA; 11392 MW; E7E1229586411A56 CRC64;

Query Match 59.6%; Score 34; DB 1; Length 108;
Best Local Similarity 63.6%; Pred. No. 8.8;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 QGDSLKTYAS 11
DB 22 QGDSLRYGDA 32
|||||

RESULT 10
PHAB SYN3 STANDARD; PRT; 161 AA.
ID PHAB SYN3
AC Q01952;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Allophycocyanin beta chain.

GN APCB OR SLR1986.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93054612; PubMed=1429645;
RA Su X., Goodman P., Bogorad L.;
RT "Excitation energy transfer from phycocyanin to chlorophyll in an
RT apha-defective mutant of Synechocystis sp. PCC 6803.";
RL J. Biol. Chem. 267:22944-22950(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hiroseawa M., Sugita M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K.,
RA Okumura S., Shimpo S., Takeuchi C., Wada T., Watanabe A.,
RA Yamada M., Yasuda M., Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
RN [3]
RP SEQUENCE OF 1-20.
RX MEDLINE=97443974; PubMed=9298645;
RA Sazuka T., Ohara O.;
RT "Towards a proteome project of cyanobacterium Synechocystis sp.
RT strain PCC6803: linking 130 protein spots with their respective
RT genes.";
RL Electrophoresis 18:1252-1258(1997).
CC -!- FUNCTION: Light-harvesting photosynthetic bile pigment-protein
CC from the phycobiliprotein complex. Allophycocyanin has a maximum
CC absorption at approximately 650 nanometers.
CC -!- SUBUNIT: Heterodimer of an alpha and a beta chain.
CC -!- PTM: Contains one covalently linked bilin chromophore.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M77135; AAA27277.1; -;
DR EMBL; D90910; BAA17875.1; -;
DR PIR; B44462; B44462.
DR HSSP; P00318; LB33.
DR InterPro; IPR006245; ApCB.
DR ProDom; IPR001659; Phycobilisome.
DR Pfam; PF00502; Phycobilisome; 1.
DR ProDom; PD000340; Phycobilisome; 1.
DR TIGRFAMs; TIGR01337; apCB; 1.
KW Phycobilisome; Electron transport; Photosynthesis; Bile pigment;
KW Methylation; Complete proteome.
FT MOD_RES 71 71 METHYLATION (BY SIMILARITY).
FT BINDING 81 81 PHYCOYANOBILIN CHROMOPHORE
FT (BY SIMILARITY).
SQ SEQUENCE 161 AA; 17215 MW; 63F01E5903BA1B83 CRC64;
Query Match 59.6%; Score 34; DB 1; Length 161;
Best Local Similarity 66.7%; Pred. No. 13;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 3 DSLKTYAS 11
Db 25 DSLKSYFAS 33
RESULT 11
XYLA_LACLA
ID XYLA_LACLA STANDARD; PRT; 439 AA.

AC O9CFG7: Q9RAV8; Q9X416; Q9X422;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Xylose isomerase (EC 5.3.1.5).
GN XYLA OR IL1509.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=210, NRRL B-4449, and IO-1;
RA Erlanson K.A., Park J.-H., Delamarre S.C., El Khal W., Kao H.-H.,
RA Basaran P., Brydges S.D., Batt C.A.;
RT "The xylose and xylan loci of Lactococcus lactis.";
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=IL1403;
RX MEDLINE=21235186; PubMed=11337471;
RA Bolotin A., Wincker P., Mauger S., Jaillon O., Malarne K.,
RA Weissenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
RT lactis ssp. lactis IL1403.";
RL Genome Res. 11:731-753(2001).
CC -!- CATALYTIC ACTIVITY: D-xylose = D-xylulose.
CC -!- COFACTOR: Binds 2 magnesium ions per subunit (By similarity).
CC -!- SUBUNIT: Homotetramer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the xylose isomerase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF092040; AAD20243.1; -;
DR EMBL; AF092041; AAD20249.1; -;
DR EMBL; AF092042; AAD20255.1; -;
DR EMBL; AE006381; AAK05607.1; -;
DR PIR; E86813; E86813.
DR HSSP; P54273; IA0D.
DR HAMAP; MF_00455; -; 1.
DR InterPro; IPR001998; Xylose isom.
DR Pfam; PF00259; Xylose isom; 1.
DR PRINTS; P000688; XYLOSISMRASE.
DR PROSITE; PS00172; XYLOSE-ISOMERASE_1; 1.
DR PROSITE; PS00173; XYLOSE-ISOMERASE_2; 1.
KW Isomerase; Pentose shunt; Xylose metabolism; Metal-binding; Magnesium;
KW Complete proteome.
FT ACT_SITE 101 101 BY SIMILARITY.
FT METAL 232 232 MAGNESIUM 1 (BY SIMILARITY).
FT METAL 268 268 MAGNESIUM 1 AND 2 (BY SIMILARITY).
FT METAL 271 271 MAGNESIUM 2 (BY SIMILARITY).
FT METAL 296 296 MAGNESIUM 1 (BY SIMILARITY).
FT METAL 307 307 MAGNESIUM 2 (BY SIMILARITY).
FT METAL 309 309 MAGNESIUM 2 (BY SIMILARITY).
FT METAL 339 339 MAGNESIUM 1 (BY SIMILARITY).
FT VARIANT 202 202 R -> M (IN STRAINS NRRL B-4449 AND IO-1).
FT VARIANT 218 218 D -> Y (IN STRAIN 210).
FT VARIANT 247 247 A -> S (IN STRAINS NRRL B-4449 AND IO-1).
FT VARIANT 275 275 T -> A (IN STRAINS NRRL B-4449 AND IO-1).
FT VARIANT 388 388 V -> S (IN STRAINS NRRL B-4449 AND IO-1).
FT VARIANT 407 407 E -> K (IN STRAIN IO-1).
FT VARIANT 416 416 H -> Y (IN STRAIN NRRL B-4449).
SQ SEQUENCE 439 AA; 49618 MW; 774A3655881B759 CRC64;
Query Match 59.6%; Score 34; DB 1; Length 439;
Best Local Similarity 60.0%; Pred. No. 40;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```

QY      2 GDSLKTYTAS 11
DB      109 GNSLKEFYAN 118

RESULT 12
ENO2_MAIZE STANDARD; PRT; 446 AA.
AC P42895;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Enolase 2 (EC 4.2.1.11) (2-phosphoglycerate dehydratase 2) (2-phospho-
DE D-glycerate hydro-lyase 2).
GN ENO2.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. B73; TISSUE=Root;
RX MEDLINE=99063764; PubMed=9847102;
RA Lal S.K., Lee C., Sachs M.M.;
RT "Differential regulation of enolase during anaerobiosis in maize.";
RL Plant Physiol. 118:1285-1293(1998).
CC -!- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = phosphoenolpyruvate +
CC H(2)O.
CC
CC -!- COFACTOR: Magnesium is required for catalysis and for stabilizing
CC the dimer (By similarity).
CC -!- PATHWAY: Glycolysis.
CC -!- SUBUNIT: Homodimer.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the enolase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U17973; AAD04187.1; -.
DR PIR; T02221; T02221.
DR HSSP; P56252; LPDZ.
DR MaizeDB; 30060; -.
DR InterPro; IPR000941; Enolase.
DR Pfam; PF00113; enolase; 1.
DR PRINTS; PR00148; ENOLASE.
DR ProDom; PD000902; Enolase; 1.
DR TIGRFAMs; TIGR01060; eno; 1.
DR PROSITE; PS00164; ENOLASE; 1.
KW Lyase; Glycolysis; Magnesium; Multigene family.
FT ACT_SITE 164 164
FT METAL 251 251
FT METAL 302 302
FT METAL 329 329
SQ SEQUENCE 446 AA; 48162 MW; DC27708CF92F6850 CRC64;

Query Match 59.6%; Score 34; DB 1; Length 446;
Best Local Similarity 70.0%; Pred. No. 41;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      2 GDSLKTYTAS 11
DB      282 GDSLKNVYKS 291

RESULT 13
ENO2_MAIZE STANDARD; PRT; 446 AA.
AC P42895;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Enolase 2 (EC 4.2.1.11) (2-phosphoglycerate dehydratase 2) (2-phospho-
DE D-glycerate hydro-lyase 2).
GN ENO2.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. B73; TISSUE=Root;
RX MEDLINE=99063764; PubMed=9847102;
RA Lal S.K., Lee C., Sachs M.M.;
RT "Differential regulation of enolase during anaerobiosis in maize.";
RL Plant Physiol. 118:1285-1293(1998).
CC -!- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = phosphoenolpyruvate +
CC H(2)O.
CC
CC -!- COFACTOR: Magnesium is required for catalysis and for stabilizing
CC the dimer (By similarity).
CC -!- PATHWAY: Glycolysis.
CC -!- SUBUNIT: Homodimer.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the enolase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U17973; AAD04187.1; -.
DR PIR; T02221; T02221.
DR HSSP; P56252; LPDZ.
DR MaizeDB; 30060; -.
DR InterPro; IPR000941; Enolase.
DR Pfam; PF00113; enolase; 1.
DR PRINTS; PR00148; ENOLASE.
DR ProDom; PD000902; Enolase; 1.
DR TIGRFAMs; TIGR01060; eno; 1.
DR PROSITE; PS00164; ENOLASE; 1.
KW Lyase; Glycolysis; Magnesium; Multigene family.
FT ACT_SITE 164 164
FT METAL 251 251
FT METAL 302 302
FT METAL 329 329
SQ SEQUENCE 446 AA; 48162 MW; DC27708CF92F6850 CRC64;

Query Match 59.6%; Score 34; DB 1; Length 446;
Best Local Similarity 70.0%; Pred. No. 41;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      2 GDSLKTYTAS 11
DB      282 GDSLKNVYKS 291

RESULT 14.
HEAD BPPZAV.
ID HEAD BPPZA STANDARD; PRT; 448 AA.
AC P07531;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Major head protein (Late protein Gp8).

```

```

GN OS Bacteriophage PZA.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
OC Phi-29-like viruses.
OX NCBI_TaxID=10757;
RN [1]
RX MEDLINE=87031573; PubMed=3095188;
RA Paces V., Vlcek C., Urbanek P.;
RT "Nucleotide sequence of the late region of Bacillus subtilis phage
RL PZA, a close relative of phi 29.";
RL Gene 44:107-114(1986).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M11813; AAA88484.1; -.
DR PIR; B24831; WMBP82.
DR InterPro; IPR003343; Big_2.
DR Pfam; PF02368; Big_2; 1.
DR SMART; SM00635; BID_2; 1.
KW Late protein.
SQ SEQUENCE 448 AA; 49754 MW; 70B6108E556BBCB2 CRC64;

Query Match 59.6%; Score 34; DB 1; Length 448;
Best Local Similarity 75.0%; Pred. No. 41;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 QGDSLKTY 8
DB 28 QGDSFKSY 35

RESULT 15
LREI_YEAST
ID LREI_YEAST STANDARD; PRT; 586 AA.
AC P25579;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Laminarase-resistance LREI.
LREI OR YCL051W OR YCL51W.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RX SEQUENCE FROM N.A.
RA Fuller L.J., Kelly A., Lewis C., McKee R.A., Pearson B.M.;
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
RN [2]
RX CHARACTERIZATION.
RX MEDLINE=97245290; PubMed=9090049;
RA Lai M.H., Silverman S.J., Gaughran J.P., Kirsch D.R.;
RT "Multiple copies of PBS2, MHP1 or LRE1 produce glucanase resistance
RT and other cell wall effects in Saccharomyces cerevisiae.";
RL Yeast 13:199-213(1997).
CC -!- FUNCTION: Not known; overexpression of both LRE1 and PBN1 confers
CC resistance to laminarinase.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X59720; CAA42393.1; -.
DR PIR; S19381; S19381.
DR Germline; 138887; -.
DR SGD; S0000556; LRE1.
DR GO; GO:0004860; P:protein kinase inhibitor activity; IDA.
DR GO; GO:0030528; P:transcription regulator activity; IDA.
DR GO; GO:0007047; P:cell wall organization and biogenesis; IGI.
DR GO; GO:0009408; P:response to heat; IDA.
SQ SEQUENCE 586 AA; 65152 MW; 7C777C5259247DC3 CRC64;

Query Match 59.6%; Score 34; DB 1; Length 586;
Best Local Similarity 66.7%; Pred. No. 54;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 DSLKTYVAS 11
DB 531 DALKTFHAS 539

RESULT 16
MY16_MOUSE
ID MY16_MOUSE STANDARD; PRT; 657 AA.
AC P17564;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Myeloid differentiation primary response protein MyD116.
GN MyD116.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RX SEQUENCE FROM N.A.
RC STRAIN=SL;
RX MEDLINE=90251472; PubMed=2339071;
RA Lord K.A., Hoffman-Liebermann B., Liebermann D.A.;
RT "Sequence of MyD116 cDNA: a novel myeloid differentiation primary
RT response gene induced by IL6.";
RL Nucleic Acids Res. 18:2823-2823(1990).
CC -!- INDUCTION: By interleukin-6.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X51829; CAA36128.1; -.
DR PIR; S10001; S10001.
DR MGD; MGI:1100516; Myd116.
KW Differentiation; Antigen; Repeat.
FT DOMAIN 283 451 4.5 X TANDEM REPEATS.
FT REPEAT 283 322 1.
FT REPEAT 323 360 2.
FT REPEAT 361 398 3.
FT REPEAT 399 436 4.
FT REPEAT 437 451 5 (INCOMPLETE).
SQ SEQUENCE 657 AA; 71840 MW; 9B217001019C38A7 CRC64;

Query Match 59.6%; Score 34; DB 1; Length 657;
Best Local Similarity 60.0%; Pred. No. 62;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 GDSLKTYVAS 11
DB 179 GETVTKYQAS 188

RESULT 17
IF38_YEAST

```

ID IP38_YEAST STANDARD; PRT; 812 AA.
AC P32497;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Eukaryotic translation initiation factor 3 93 kDa subunit (eif3 p93)
DE (Nuclear transport protein NIP1).
GN NIP1 OR YMR309C OR YMR924.01C OR YMR952.11C.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93066237; PubMed=1332047;
RA Gu Z., Moersschell R.P., Sherman F., Goldfarb D.S.;
RT "NIP1, a gene required for nuclear transport in yeast."
RL Proc. Natl. Acad. Sci. U.S.A. 89:10355-10359(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX MEDLINE=97313268; PubMed=9169872;
RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
RA Jagels K., Lye G., Moule S., Odell C., Pearson D., Rajandream M.A.,
RA Rice P., Skelton J., Walsh S., Whitehead S., Barrell B.G.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XIII."
RL Nature 387:90-93(1997).
CC -!- FUNCTION: Binds to the 40S ribosome and promotes the binding of methionyl-tRNAI and mRNA.
CC -!- FUNCTION: INVOLVED IN TRANSPORT OF PROTEINS TO THE NUCLEUS;
CC -!- SUBUNIT: eif-3 is composed of up to 8 different subunits.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic; mainly.
CC -!- SIMILARITY: Contains 1 PFI domain.
CC -!- SIMILARITY: BELONGS TO THE EIF38 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L02899; -; NOT ANNOTATED_CDS.
DR EMBL; Z54141; CA90827.1; -;
DR EMBL; Z49212; CA989142.1; -;
DR PIR; A46417; A46417.
DR GeneOnline; 142988; -;
DR SGD; S0004926; NIP1.
DR GO; GO:0005737; Cytoplasm; IDA.
DR GO; GO:0005852; C:eukaryotic translation initiation factor 3 . . .; IDA.
DR GO; GO:0003743; F:translation initiation factor activity; IDA.
DR GO; GO:0006413; P:translational initiation; IDA.
DR InterPro; IPR008905; eif3c_N.
DR InterPro; IPR000717; PFI.
DR Pfam; PF05470; eif3c_N; 1.
DR Pfam; PF01399; PFI; 1.
DR SMART; SM00088; PINT; 1.
KW Initiation factor; Protein biosynthesis; Transport; Protein transport.
FT DOMAIN 15 147 ASP/GLU-RICH (ACIDIC).
FT DOMAIN 15 36 SER-RICH.
FT CONFLICT 111 111 V -> D (IN REF. 1).
FT CONFLICT 583 583 Q -> H (IN REF. 1).
FT CONFLICT 641 641 K -> N (IN REF. 1).
FT CONFLICT 643 643 K -> N (IN REF. 1).
SQ SEQUENCE 812 AA; 93204 MW; EE05097C44C45A4C CRC64;
Query Match 59.6%; Score 34; DB 1; Length 812;
Best Local Similarity 63.6%; Pred. No. 77;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
Db 718 QVESLKTYPFS 728
RESULT 18
HBB1 TRICR
ID HBB1 TRICR STANDARD; PRT; 145 AA.
AC P10785;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hemoglobin beta-1 chain (Major).
GN HBB1.
OS Triturus cristatus (Great crested newt) (Warty newt).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Caudata; Salamandroidea; Salamandridae;
OC Triturus.
OX NCBI_TaxID=8323;
RN [1]
RP SEQUENCE.
RX MEDLINE=89207117; PubMed=3242554;
RA Kleinschmidt I., Sgouras J.G., Braunitzer G.;
RT "The first sequenced normal hemoglobin lacking histidine in position 146 of the beta-chains. The primary structures of the major and minor hemoglobin components of the great crested newt (Triturus cristatus, Urodela, Amphibia)."
RL Urodel. Chem. Hoppe-Seyler 369:1343-1360(1988).
CC -!- FUNCTION: Involved in oxygen transport from the lung to the various peripheral tissues.
CC -!- SUBUNIT: MAJOR HEMOGLOBIN IS A TETRAMER OF TWO ALPHA-1 CHAINS AND TWO BETA-1 CHAINS.
CC -!- TISSUE SPECIFICITY: Red blood cells.
CC -!- SIMILARITY: Belongs to the globin family.
DR PIR; S02025; S02025.
DR HSP; P02070; 1FSX.
DR InterPro; IPR002337; Beta haem.
DR InterPro; IPR000971; Globin.
DR Pfam; PF00042; globin; 1.
DR PRINTS; PR00814; BETAHAEM.
DR PROSITE; PS01033; GLOBIN; 1.
KW Heme; Oxygen transport; Transport; Erythrocyte.
FT METAL 62 62 IRON (HEME DISTAL LIGAND).
FT METAL 91 91 IRON (HEME PROXIMAL LIGAND).
SQ SEQUENCE 145 AA; 16003 MW; BEF06F441B42BA80 CRC64;
Query Match 57.9%; Score 33; DB 1; Length 145;
Best Local Similarity 75.0%; Pred. No. 19;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 DSLKTYVA 10
Db 78 DDLKAYYA 85
RESULT 19
PHAB SYNY4
ID PHAB SYNY4 STANDARD; PRT; 161 AA.
AC Q02924;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-JUL-1993 (Rel. 26, Last annotation update)
DE Allophycocyanin beta chain.
GN APCB.
OS Synechocystis sp. (strain PCC 6714).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1147;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93222481; PubMed=8467079;
RA Dimagno L.M., Haselkorn R.;
RT "Isolation and characterization of the genes encoding allophycocyanin

```

RT subunits and two linker proteins from Synechocystis 6714."
RL Plant Mol. Biol. 21:835-846(1993).
CC -!- FUNCTION: Light-harvesting photosynthetic bile pigment-protein
CC from the phycobiliprotein complex. Allophycocyanin has a maximum
CC absorption at approximately 650 nanometers.
CC -!- SUBUNIT: Heterodimer of an alpha and a beta chain.
CC -!- PTM: Contains one covalently linked bilin chromophore.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L02308; AAA69683.1; -
DR PIR; S33624; S33624.
DR HSSP; P00316; 1B33.
DR InterPro; IPR006245; ApCB.
DR InterPro; IPR001659; Phycobilisome.
DR Pfam; PF00502; Phycobilisome; 1.
DR ProDom; PD000340; Phycobilisome; 1.
DR TIGRfam; TIGR01337; apCB; 1.
DR Phycobilisome; Electron transport; Photosynthesis; Bile pigment;
KW Methylation.
KW MOD_RES 71 71 METHYLATION (BY SIMILARITY).
FT BINDING 81 81 PHYCOCYANOBILIN CHROMOPHORE.
FT SEQUENCE 161 AA; 17242 MW; 63F0047008630A73 CRC64;
SQ
Query Match 57.9%; Score 33; DB 1; Length 161;
Best Local Similarity 66.7%; Pred. No. 21;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 DSLKTYVAS 11
DB 25 DLKXNFAS 33

RESULT 20
YQAC BACSU STANDARD; PRT; 178 AA.
AC P45900;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein yqac precursor.
GN YQAC OR BSU26370.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / JH642;
RX MEDLINE=95219086; PubMed=7704261;
RA Takemaru K.-I., Mizuno M., Sato T., Takeuchi M., Kobayashi Y.;
RT "Complete nucleotide sequence of a skin element excised by DNA
RT rearrangement during sporulation in Bacillus subtilis."
RL Microbiology 141:323-327(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / JH642;
RX MEDLINE=97124195; PubMed=8969508;
RA Mizuno M., Masuda S., Takemaru K.-I., Hosono S., Sato T., Takeuchi M.,
RA Kobayashi Y.;
RT "Systematic sequencing of the 283 kb 210 degrees-232 degrees region of
RT the Bacillus subtilis genome containing the skin element and many
RT sporulation genes."
RL Microbiology 142:3103-3111(1996).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;

```

```

RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriess R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Davine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerz-Blanchard M., Klein C.,
RA Kobayashi Y., Koeter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapluda A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauei C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Fujic P., Fumelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadale Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tostato V., Uchiyama S., Vandenbol M., Vanier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis";
RL Nature 390:249-256(1997).
RN [4]
RP IDENTIFICATION.
RX MEDLINE=96084975; PubMed=7489895;
RA Medigue C., Moszer I., Viari A., Danchin A.;
RT "Analysis of a Bacillus subtilis genome fragment using a co-operative
RT computer system prototype.";
RL Gene 165:GC37-GC51(1995).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D32216; BAA06916.1; -
DR EMBL; D84432; BAA12377.1; -
DR EMBL; Z99117; CAB14578.1; -
DR PIR; B69944; B69944.
DR Subtilist; BG11254; yqac.
KW Hypothetical protein; Signal; Complete proteome.
FT SIGNAL 1 19
FT CHAIN 20 178 HYPOTHEICAL PROTEIN YQAC.
SQ SEQUENCE 178 AA; 20702 MW; DD2DE09D65CF882E CRC64;
Query Match 57.9%; Score 33; DB 1; Length 178;
Best Local Similarity 62.5%; Pred. No. 24;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 GDSLKTYTY 9
DB 75 GDNLSYF 82

RESULT 21
KSGA_PSEAE STANDARD; PRT; 268 AA.
AC Q915U5;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)

```

DE Dimethyladenosine transferase (EC 2.1.1.1.-) (S-adenosylmethionine-6-N',
DE N'-adenosyl(rRNA) dimethyltransferase) (16S rRNA dimethylase) (High
DE level kasugamycin resistance protein ksgA) (Kasugamycin
DE dimethyltransferase).
GN KSGA OR PA0592
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen";
RL Nature 406:959-964(2000).
CC -!- FUNCTION: Specifically dimethylates two adjacent adenosines in the
CC loop of a conserved hairpin near the 3' end of 16S rRNA in the 30S
CC particle. Its inactivation leads to kasugamycin resistance (by
CC similarity).
CC -!- SIMILARITY: Belongs to the rRNA adenine N-6-methyltransferase
CC family. KsgA subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB004495; H83571.
DR PIR; H83571; H83571.
DR HAMAP; MF_00607; -; 1.
DR InterPro; IPR0001737; RNA A dimeth.
DR InterPro; IPR0000051; SAM Bind.
DR Pfam; PF00398; RnaAD; 1.
DR SMART; SM00650; rADC; 1.
DR TIGRFAMs; TIGR00755; ksgA; 1.
DR PROSITE; PS01131; RNA A DIMETH; 1.
KW rRNA processing; Transferase; Methyltransferase;
KW Antibiotic resistance; Complete proteome.
SQ SEQUENCE 268 AA; 30058 MW; A2A596FF4341193A CRC64;

Query Match 57.9%; Score 33; DB 1; Length 268;
Best Local Similarity 63.6%; Pred. No. 37;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
Db |||:|:|:|:
87 QGDALKDFAS 97

RESULT 22
ID RECX LACLA STANDARD; PRT; 270 AA.
AC Q9CDN7;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Regulatory protein recX.
GN RECX OR LL2182.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=IL1403;
RX MEDLINE=21235186; PubMed=11337471;
RA Bolotin A., Wincker P., Mauger S., Jallou O., Malarne K.,
RA Weissenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
RT lactis ssp. lactis IL1403";
RL Genome Res. 11:731-753(2001).
CC -!- FUNCTION: Modulates recA activity (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -!- SIMILARITY: Belongs to the recX family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF006447; AAK06280.1; -.
DR PIR; F86897; F86897.
DR HAMAP; MF_01114; -; 1.
DR InterPro; IPR003783; RecX.
DR Pfam; PF02631; RecX; 1.
KW Complete proteome.
SQ SEQUENCE 270 AA; 31270 MW; 1968460506330C4D CRC64;

Query Match 57.9%; Score 33; DB 1; Length 270;
Best Local Similarity 63.6%; Pred. No. 38;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
Db |||:|:|:|:
61 QGKSLALYIS 71

RESULT 23
ID THTR CHICK STANDARD; PRT; 289 AA.
AC P25324;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Thiosulfate sulfurtransferase (EC 2.8.1.1) (Rhodanese).
GN TST.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE
RC TISSUE=Liver;
RX MEDLINE=91113289; PubMed=2275748;
RA Kohanski R.A., Heinrikson R.L.;
RT "Primary structure of avian hepatic rhodanese";
RL J. Protein Chem. 9:369-377(1990).
CC -!- FUNCTION: FORMATION OF IRON-SULFUR COMPLEXES AND CYANIDE
CC DETOXIFICATION.
CC -!- CATALYTIC ACTIVITY: Thiosulfate + cyanide = sulfite + thiocyanate.
CC -!- SUBUNIT: Monomer.
CC -!- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC -!- TISSUE SPECIFICITY: Expressed in numerous tissues.
CC -!- DOMAIN: THE STRUCTURE CONSISTS OF 2 DOMAINS OF VERY SIMILAR
CC CONFORMATION, SUGGESTING A COMMON EVOLUTIONARY ORIGIN. HOWEVER,
CC THE SEQUENCES OF THE 2 DOMAINS ARE VERY DIFFERENT.
CC -!- SIMILARITY: Contains 2 rhodanese domains.
DR PIR; A37209; A37209.
DR HSP; P00586; 1RHS.
DR InterPro; IPR001763; Rhodanese-like.
DR InterPro; IPR001307; Rhodanese.
DR Pfam; PF00581; Rhodanese; 2.
DR SMART; SM00450; RHOD; 2.

DR PROSITE; PS00380; RHODANESE 1; 1.
 DR PROSITE; PS00683; RHODANESE 2; 1.
 DR PROSITE; PS0206; RHODANESE 3; 2.
 KW Transferrase; Mitochondrion; Repeat.
 FT DOMAIN 24 142 RHODANESE 1.
 FT DOMAIN 143 158 HINGE.
 FT DOMAIN 172 284
 FT ACT_SITE 186 186
 FT ACT_SITE 244 244
 FT ACT_SITE 245 245
 FT ACT_SITE 246 246
 FT ACT_SITE 246 246
 SQ SEQUENCE 289 AA; 32286 MW; 8BFCF671DE0B2BA4 CRC64;
 Query Match 57.9%; Score 33; DB 1; Length 289;
 Best Local Similarity 66.7%; Pred. No. 40;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 2 GDSLKTYYA 10
 DB 100 GDELGTFFA 108
 ID PRTN TRIAL STANDARD; PRT; 387 AA.
 AC P23653;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Proteinase R precursor (SC 3.4.21.-).
 GN PROR.
 OS Triticarium album.
 OC Eukaryota; Fungi; Ascomycota; mitosporic Ascomycota; Triticarium.
 OX NCBI_TaxID=5556;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 22563 / Limber;
 RX MEDLINE=91171891; PubMed=2077361;
 RA Samal B.B., Karan B., Boone T.C., Osslund T.D., Chen K.K.,
 RA Stabinsky Y.;
 RT "Isolation and characterization of the gene encoding a novel,
 RT thermostable serine proteinase from the mould Triticarium album
 RT Limber.";
 RL Mol. Microbiol. 4:1789-1792(1990).
 CC -!- FUNCTION: Thermostable serine proteinase.
 CC -!- COFACTOR: Binds 2 calcium ions per subunit (By similarity).
 CC -!- SIMILARITY: Belongs to peptidase family S8.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X56116; CAA39584.1; -.
 DR PIR; S11985; S11985.
 DR HSSP; P06873; 2PRK.
 DR MEROPS; S08.054; -.
 DR InterPro; IPR000209; Peptidase S8.
 DR InterPro; IPR009020; Protease Inhib.
 DR Pfam; PF00082; Peptidase S8; 1.
 DR PRINTS; PR00723; SUBTILISIN.
 DR PROSITE; PS00136; SUBTILASE ASP; 1.
 DR PROSITE; PS00137; SUBTILASE HIS; 1.
 DR PROSITE; PS00138; SUBTILASE SER; 1.
 DR Hydrolase; Serine protease; Zymogen; Calcium-binding; Signal.
 KW SIGNAL 1 21 POTENTIAL.
 FT PROPEP 22 108

FT CHAIN 109 387
 FT ACT_SITE 147 147
 FT ACT_SITE 177 177
 FT ACT_SITE 332 332
 FT METAL 124 124
 FT METAL 308 308
 FT METAL 368 368
 FT DISULFID 142 231
 FT DISULFID 286 357
 SQ SEQUENCE 387 AA; 40926 MW; 66E9B953FADDA188 CRC64;
 Query Match 57.9%; Score 33; DB 1; Length 387;
 Best Local Similarity 85.7%; Pred. No. 55;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 5 LKTYVAS 11
 DB 164 VKTYVAS 170
 ID ENO1 MAIZE STANDARD; PRT; 446 AA.
 AC P26301;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Enolase 1 (EC 4.2.1.11) (2-phosphoglycerate dehydratase 1) (2-phospho-
 DE D-glycerate hydro-lyase 1).
 GN ENO1 OR PGH1.
 OS Zea mays (Maize).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
 OX NCBI_TaxID=4577;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Berkeley Fast; TISSUE=Root;
 RX MEDLINE=91316216; PubMed=1859865;
 RA Lal S.K., Johnson S., Conway T., Kelley P.M.;
 RA "Characterization of a maize cDNA that complements an enolase-
 RA deficient mutant of Escherichia coli.";
 RL Plant Mol. Biol. 16:787-795(1991).
 CC -!- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = phosphoenolpyruvate +
 CC H₂O.
 CC -!- COFACTOR: Magnesium is required for catalysis and for stabilizing
 CC the dimer (By similarity).
 CC -!- PATHWAY: Glycolysis.
 CC -!- SUBUNIT: Homodimer.
 CC -!- SIMILARITY: Belongs to the enolase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X55981; CAA39454.1; -.
 DR PIR; S16257; S16257.
 DR HSSP; P56252; 1PDZ.
 DR MaizeDB; 30060; -.
 DR InterPro; IPR000941; Enolase.
 DR Pfam; PF00113; enolase; 1.
 DR Pfam; PF03952; enolase N; 1.
 DR PRINTS; PR00148; ENOLASE.
 DR ProDom; PD000902; Enolase; 1.
 DR TIGRFAMs; TIGR01060; eno; 1.
 DR PROSITE; PS00164; ENOLASE; 1.
 KW Lyase; Glycolysis; Magnesium; Multigene family.

FT ACT SITE 164 164 BY SIMILARITY.
 FT METAL 251 251 MAGNESIUM (BY SIMILARITY).
 FT METAL 302 302 MAGNESIUM (BY SIMILARITY).
 FT METAL 329 329 MAGNESIUM (BY SIMILARITY).
 SQ SEQUENCE 446 AA; 48063 MW; 6266C48914F35198 CRC64;

 Query Match 57.9%; Score 33; DB 1; Length 446;
 Best Local Similarity 70.0%; Pred. No. 65;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

 QY 2 GDSLKTYTAS 11
 Db 282 GDSLKDLYKS 291

Search completed: September 24, 2004, 01:55:13
 Job time : 14.1321 secs

This Page Blank (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 01:47:17 ; Search time 52.717 Seconds
(without alignments)
65.836 Million cell updates/sec

Title: US-10-088-639a-2_COPY_23_33
Perfect score: 57
Sequence: 1 QGSLKTYAS 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database :

- SPTREMBL_25:*
- 1: sp archaea:*
 - 2: sp bacteria:*
 - 3: sp fungi:*
 - 4: sp human:*
 - 5: sp invertebrate:*
 - 6: sp mammal:*
 - 7: sp mhc:*
 - 8: sp organelle:*
 - 9: sp phage:*
 - 10: sp plant:*
 - 11: sp rodent:*
 - 12: sp virus:*
 - 13: sp vertebrate:*
 - 14: sp unclassified:*
 - 15: sp_rvirus:*
 - 16: sp_bacteriap:*
 - 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	50	87.7	107	4 Q9NSD6	Q9nsd6 homo sapien
2	46	80.7	81	4 Q7Z2E8	Q7z2e8 homo sapien
3	44	77.2	161	5 O96405	O96405 plasmodium
4	44	77.2	192	5 O8WPY5	O8wp5 plasmodium
5	44	77.2	2135	5 O61077	O61077 plasmodium
6	40	70.2	55	5 Q9TVZ4	Q9tvz4 plasmodium
7	39	68.4	54	5 Q9U7L1	Q9u7l1 plasmodium
8	39	68.4	54	5 Q9U7K5	Q9u7k5 plasmodium
9	39	68.4	60	5 Q9BJJ1	Q9bjj1 plasmodium
10	39	68.4	72	5 Q9ND13	Q9nd13 plasmodium
11	39	68.4	74	5 Q9ND11	Q9nd11 plasmodium
12	39	68.4	179	5 Q9U7J9	Q9u7j9 plasmodium
13	39	68.4	267	5 O61076	O61076 plasmodium
14	39	68.4	516	5 Q9VU83	Q9vu83 drosophila
15	39	68.4	542	5 Q9VU82	Q9vu82 drosophila
16	39	68.4	609	5 Q9VU81	Q9vu81 drosophila

17	39	68.4	1086	12 Q8JK71	Q8jk71 heliotis vi
18	39	68.4	1729	5 Q25734	Q25734 plasmodium
19	39	68.4	2162	5 Q8IEV1	Q8iev1 plasmodium
20	39	68.4	2163	5 Q9NFB6	Q9nfb6 plasmodium
21	39	68.4	2207	5 Q8I495	Q8i495 plasmodium
22	39	68.4	2924	5 Q25733	Q25733 plasmodium
23	39	68.4	9271	5 Q8IES7	Q8ies7 plasmodium
24	38	66.7	54	5 Q9U7K6	Q9u7k6 plasmodium
25	38	66.7	56	5 Q9VJ78	Q9vj78 drosophila
26	38	66.7	90	16 Q8NK25	Q8nk25 xanthomona
27	38	66.7	110	9 Q21984	Q21984 streptococ
28	38	66.7	327	5 Q17032	Q17032 anopheles g
29	38	66.7	347	5 Q8I2D6	Q8i2d6 plasmodium
30	38	66.7	429	16 Q25470	Q25470 helicobacte
31	38	66.7	429	16 Q9ZL62	O9z162 helicobacte
32	38	66.7	992	5 Q8ILF1	Q8ilf1 plasmodium
33	38	66.7	1697	12 Q8UJG7	Q8ujg7 norwalk-lik
34	37	64.9	228	2 Q7WU96	Q7wu96 cycloclasti
35	37	64.9	288	11 Q8R0D2	Q8r0d2 mus musculu
36	37	64.9	297	11 Q80UW2	Q80uw2 mus musculu
37	37	64.9	313	16 Q8R7K6	Q8r7k6 thermaoer
38	37	64.9	379	5 Q9V554	Q9v554 drosophila
39	37	64.9	737	3 Q9HDE9	Q9hde9 metarizium
40	37	64.9	1077	1 Q977Q4	Q977q4 methanosarc
41	37	64.9	1077	1 Q977X0	Q977x0 methanosarc
42	37	64.9	1077	1 Q977X1	Q977x1 methanosarc
43	37	64.9	1095	17 Q8PXS2	Q8pxs2 methanosarc
44	37	64.9	1340	5 Q8ICF1	Q8icf1 plasmodium
45	36.5	64.0	1042	16 Q8A7I4	Q8a7i4 bacteroides
46	36	63.2	208	16 Q82W52	Q82w52 nitrosomona
47	36	63.2	233	16 Q81CU1	Q81cu1 bacillus ce
48	36	63.2	283	16 Q8CX77	Q8cx77 oceanobacil
49	36	63.2	306	5 Q8IRS6	Q8irs6 drosophila
50	36	63.2	306	5 Q95SR5	Q95sr5 drosophila
51	36	63.2	369	3 Q8UJQ7	Q8ujq7 gibberella
52	36	63.2	452	4 Q8TEY0	Q8tey0 homo sapien
53	36	63.2	516	16 Q82NP2	Q82np2 streptomyce
54	36	63.2	545	16 Q89AX4	Q89ax4 buchnera ap
55	36	63.2	721	16 Q8A3J0	Q8a3j0 bacteroides
56	36	63.2	1680	12 Q9YQ23	Q9yq23 bovine cali
57	35	61.4	26	4 Q16021	Q16021 homo sapien
58	35	61.4	53	5 Q9U7F9	Q9u7f9 plasmodium
59	35	61.4	71	5 Q9ND30	Q9nd30 plasmodium
60	35	61.4	71	5 Q9ND10	Q9nd10 plasmodium
61	35	61.4	72	5 Q9ND31	Q9nd31 plasmodium
62	35	61.4	95	10 Q42271	Q42271 arabidopsis
63	35	61.4	111	9 Q48383	Q48383 streptococ
64	35	61.4	174	16 Q8A7S6	Q8a7s6 bacteroides
65	35	61.4	189	5 Q9U7L3	Q9u7l3 plasmodium
66	35	61.4	197	16 Q914E7	Q914e7 pseudomonas
67	35	61.4	269	17 Q9HN25	Q9hn25 halobacteri
68	35	61.4	273	16 Q893C4	Q893c4 clostridium
69	35	61.4	313	17 Q9YFA4	Q9yfa4 aeropyrum p
70	35	61.4	367	5 Q17995	Q17995 caenorhabdi
71	35	61.4	405	10 Q9SEI1	Q9sei1 arabidopsis
72	35	61.4	407	16 Q8AA63	Q8aa63 bacteroides
73	35	61.4	408	16 Q8YSM6	Q8ysm6 anabaena sp
74	35	61.4	418	10 Q9SCD6	Q9scd6 fagus sylv
75	35	61.4	419	10 Q9CSU3	Q9csu3 arabidopsis
76	35	61.4	434	16 Q819A1	Q819a1 bacillus ce
77	35	61.4	453	2 Q51716	Q51716 pseudomonas
78	35	61.4	561	10 Q8RU48	Q8ru48 oryza sativ
79	35	61.4	563	10 Q8LSK5	Q8lsk5 gossypium h
80	35	61.4	580	3 P87258	P87258 trichoderma
81	35	61.4	580	3 Q8NIN7	Q8nin7 trichoderma
82	35	61.4	652	16 Q8DT49	Q8dt49 streptococ
83	35	61.4	738	2 Q8KKH8	Q8kkh8 paenibacill
84	35	61.4	921	16 Q8UJW1	Q8ujw1 agrobacteri
85	35	61.4	1505	16 Q8DE36	Q8de36 vibrio vuln
86	35	61.4	1886	16 Q82XF9	Q82xf9 nitrosomona
87	35	61.4	2228	5 Q60991	Q60991 plasmodium
88	34	59.6	112	17 Q97BS6	Q97bs6 thermoplasma
89	34	59.6	184	16 Q8F688	Q8f688 leptospira

```
90 34 59.6 189 16 024901
91 34 59.6 189 16 032N04
92 34 59.6 212 11 08C579
93 34 59.6 212 11 08C579
94 34 59.6 218 16 067537
95 34 59.6 219 9 08W643
96 34 59.6 275 16 07TV69
97 34 59.6 278 16 0891J3
98 34 59.6 280 16 07XAS6
99 34 59.6 290 16 082A0
100 34 59.6 300 11 08VD40
101 34 59.6 303 10 08LMD5
102 34 59.6 319 3 09HEC8
103 34 59.6 384 16 083ED3
104 34 59.6 428 10 08LM12
105 34 59.6 433 16 088CR5
106 34 59.6 446 10 07XBE4
107 34 59.6 464 16 082HA9
108 34 59.6 466 12 08JXJ9
109 34 59.6 466 12 08JXK0
110 34 59.6 491 2 09ZNH8
111 34 59.6 494 2 085527
112 34 59.6 529 17 09QR8
113 34 59.6 561 16 09X857
114 34 59.6 568 16 08XK20
115 34 59.6 578 3 078738
116 34 59.6 579 10 09LH59
117 34 59.6 604 3 093847
118 34 59.6 695 17 098026
119 34 59.6 697 5 09VEP8
120 34 59.6 723 16 082BJ0
121 34 59.6 743 16 08RG20
122 34 59.6 908 5 08WMF4
123 34 59.6 987 16 08G4D8
124 34 59.6 1553 10 08L8G9
125 33 57.9 59 5 09BJI8
126 33 57.9 59 5 09BJG1
127 33 57.9 60 5 09BJI4
128 33 57.9 62 5 09BJJ5
129 33 57.9 62 5 09BUH1
130 33 57.9 62 5 09BJJ4
131 33 57.9 63 5 09BJI9
132 33 57.9 64 5 09BJI6
133 33 57.9 72 5 09ND28
134 33 57.9 105 2 086668
135 33 57.9 131 5 09V8Y8
136 33 57.9 140 10 084XK8
137 33 57.9 175 10 084XK9
138 33 57.9 177 10 09SXY9
139 33 57.9 187 17 09T0N5
140 33 57.9 215 16 09CJT9
141 33 57.9 222 16 07VF65
142 33 57.9 235 16 09K1B3
143 33 57.9 235 16 09JX89
144 33 57.9 239 16 082MM1
145 33 57.9 260 16 07L871
146 33 57.9 260 16 07TWA1
147 33 57.9 274 16 08G659
148 33 57.9 278 12 098660
149 33 57.9 296 11 0921X8
150 33 57.9 297 10 08LBW1
```

ALIGNMENTS

```
RESULT 1
Q9NSD6 PRELIMINARY; PRT; 107 AA.
AC Q9NSD6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
```

```
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymphocytes;
RA Hohmann A.;
RT "Autoimmunity.";
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; LA3092; AAA69746.2; -.
DR HSP; P01709; 2MCG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003596; IG_V.
DR Pfam; PF00047; ig; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00835; IG_LIKE; 1.
FT NON_TER 1
FT NON_TER 107
SQ SEQUENCE 107 AA; 11306 MW; A2B04B37187A5F00 CRC64;

Query Match 87.7%; Score 50; DB 4; Length 107;
Best Local Similarity 81.8%; Pred. No. 0.064;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYAS 11
DB 21 QGDSLRSFYAS 31
|||||:||||
|||||:||||

RESULT 2
Q7Z2E8 PRELIMINARY; PRT; 81 AA.
AC Q7Z2E8;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Rearranged V131 segment (Rearranged V131 gene segment)
DE (Fragment).
GN VL3L.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Hodgkin lymphoma, and Mantle cell lymphoma;
RA Tinguely M., Rosenquist R., Sundstroem C., Amini R.M., Kuppers R.,
RA Hansmann M.L., Brauning A.;
RT "Analysis of a clonally related mantle cell and Hodgkin lymphoma
RT indicates Epstein-Barr virus infection of a Hodgkin/Reed-Sternberg
RT cell precursor in a germinal center.";
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ564423; CAD92030.1; -.
DR EMBL; AJ564424; CAD92031.1; -.
FT NON_TER 1
FT NON_TER 81
SQ SEQUENCE 81 AA; 8702 MW; CFF1D466B794C9F CRC64;

Query Match 80.7%; Score 46; DB 4; Length 81;
Best Local Similarity 72.7%; Pred. No. 0.29;
Matches 8; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYAS 11
DB 4 QGDSLRSFYAS 14
|||||:||||
|||||:||||

RESULT 3
Q96405 PRELIMINARY; PRT; 161 AA.
AC Q96405
```

```

DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE Erythrocyte membrane protein 1 (Fragment).
GN FCR3S1.2VAR9.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RA Chen Q., Wahlgren M.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF039281; AAD02164.1; -.
FT NON_TER 1 1
FT NON_TER 161 161
SQ SEQUENCE 161 AA; 19005 MW; 164F7D8D9E1AE384 CRC64;

Query Match 77.2%; Score 44; DB 5; Length 161;
Best Local Similarity 77.8%; Pred. No. 1.5;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTTY 9
Db 44 EGDSEIKTY 52

RESULT 4
ID Q8WPV5 PRELIMINARY; PRT; 192 AA.
AC Q8WPV5;
DT 01-MAY-2002 (TrEMBLrel. 20, Created)
DT 01-MAY-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
DE Plasmodium falciparum.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RA Fernandez V., Chen Q., Sundstroem A., Scherf A., Hagblom P.,
RA Wahlgren M.;
RT "Mosaic-like transcription of var genes in single Plasmodium
RT falciparum parasites.";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ429510; CAD22362.1; -.
KW Hypothetical protein.
FT NON_TER 1 1
FT NON_TER 192 192
SQ SEQUENCE 192 AA; 22196 MW; 022557CE5559DC4C CRC64;

Query Match 77.2%; Score 44; DB 5; Length 192;
Best Local Similarity 77.8%; Pred. No. 1.8;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTTY 9
Db 43 EGDSEIKTY 51

RESULT 5
ID Q61077 PRELIMINARY; PRT; 2135 AA.
AC Q61077;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Variant-specific surface protein (Fragment).
GN VARPH17.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE=20183858; PubMed=10717306;
RA Voss T.S., Thompson J.K., Waterkeyn J., Felger I., Weiss N.,
RA Cowan A.F., Beck H.P.;
RT "Genomic distribution and functional characterisation of two distinct
RT and conserved Plasmodium falciparum var gene 5' flanking sequences.";
RL Mol. Biochem. Parasitol. 107:103-115(2000).
DR EMBL; AF050740; AAC05220.1; -.
DR PIR; T14602; T14602.
DR GO; GO:0005539; F:glycosaminoglycan binding; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR004258; PFEMP.
DR Pfam; PF03011; PFEMP; 2.
FT NON_TER 2135 2135
SQ SEQUENCE 2135 AA; 242712 MW; C42CEB0C32F4A36F CRC64;

Query Match 77.2%; Score 44; DB 5; Length 2135;
Best Local Similarity 77.8%; Pred. No. 24;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTTY 9
Db 160 EGDSEIKTY 168

RESULT 6
ID Q9TVZ4 PRELIMINARY; PRT; 55 AA.
AC Q9TVZ4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Erythrocyte membrane protein 1 SD126F (Fragment).
GN VAR.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RA Ward C.P., Clotney G.T., Dorris M., Ji D.-D., Arnot D.E.;
RT "Analysis of Plasmodium falciparum PfEMP-1/var genes suggests that
RT recombination rearranges constrained sequences.";
RL Mol. Biochem. Parasitol. 102:167-177(1999).
DR EMBL; AF127315; AAD52795.1; -.
DR EMBL; AF127278; AAD52758.1; -.
FT NON_TER 1 1
FT NON_TER 55 55
SQ SEQUENCE 55 AA; 6259 MW; 4DDB2C769853F649 CRC64;

Query Match 70.2%; Score 40; DB 5; Length 55;
Best Local Similarity 66.7%; Pred. No. 2.9;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTTY 9
Db 31 EGDSEIKTY 39

RESULT 7
ID Q9U7L1 PRELIMINARY; PRT; 54 AA.
AC Q9U7L1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Erythrocyte membrane protein 1 SD102B (Fragment).
GN VAR.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99404835; PubMed=10477185;

```

```

RA Ward C.P., Clotvey G.T., Dorris M., Ji D.D., Arnot D.E.;
RT "Analysis of Plasmodium falciparum PfEMP-1/var genes suggests that
RL recombination rearranges constrained sequences.";
RL Mol. Biochem. Parasitol. 102:167-177(1999).
DR EMBL; AF127275; AAD52755.1; -.
FT NON TER 1 1
FT NON TER 54 54
SQ SEQUENCE 54 AA; 6050 MW; 94F71D9C1FAAD823 CRC64;

Query Match 68.4%; Score 39; DB 5; Length 54;
Best Local Similarity 66.7%; Pred. No. 4.5;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTTY 9
Db 28 EGDSEIKTHY 36

RESULT 8
Q9U7K5 PRELIMINARY; PRT; 54 AA.
ID Q9U7K5
AC Q9U7K5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Erythrocyte membrane protein 1 SD102H (Fragment).
GN VAR.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]_TaxID=5833;
RP SEQUENCE FROM N.A.
RX MEDLINE=99404835; PubMed=10477185;
RA Ward C.P., Clotvey G.T., Dorris M., Ji D.D., Arnot D.E.;
RT "Analysis of Plasmodium falciparum PfEMP-1/var genes suggests that
RL recombination rearranges constrained sequences.";
RL Mol. Biochem. Parasitol. 102:167-177(1999).
DR EMBL; AF127282; AAD52762.1; -.
FT NON TER 1 1
FT NON TER 54 54
SQ SEQUENCE 54 AA; 5978 MW; 422FFD9C1FAADA81 CRC64;

Query Match 68.4%; Score 39; DB 5; Length 54;
Best Local Similarity 66.7%; Pred. No. 4.5;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTTY 9
Db 28 EGDSEIKTHY 36

RESULT 9
Q9BJJ1 PRELIMINARY; PRT; 60 AA.
ID Q9BJJ1
AC Q9BJJ1;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Variant surface protein (Fragment).
GN VAR.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]_TaxID=5833;
RP SEQUENCE FROM N.A.
RX STRAIN=257;
RA Noqueira P.A., Wunderlich G., Tada M.S., Costa J.D.N., Menezes M.J.,
RT "Plasmodium falciparum: repertoire of expressed var genes and adhesion
RT properties to endothelial receptors of clinical isolates from patients
RT in Rondonia (Brazilian western Amazon region).";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF333287; AAK19558.1; -.

```

```

FT NON TER 1 1
FT NON TER 60 60
SQ SEQUENCE 60 AA; 7010 MW; FF173327A53AC20A CRC64;

Query Match 68.4%; Score 39; DB 5; Length 60;
Best Local Similarity 66.7%; Pred. No. 5.1;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTTY 9
Db 34 EGDSEIKTHY 42

RESULT 10
Q9ND13 PRELIMINARY; PRT; 72 AA.
ID Q9ND13
AC Q9ND13;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Variant surface protein (Fragment).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]_TaxID=5833;
RP SEQUENCE FROM N.A.
RX STRAIN=WT09;
RX MEDLINE=20372615; PubMed=10910718;
RA Kirchgatter K., Mosbach R., del Portillo H.A.;
RT "Plasmodium falciparum: DBL-1 var sequence analysis in field isolates
RL from central Brazil.";
RL Exp. Parasitol. 95:154-157(2000).
DR EMBL; AF172801; AAF89791.1; -.
FT NON TER 1 1
FT NON TER 72 72
SQ SEQUENCE 72 AA; 8198 MW; 715D9AD0DE2D5785 CRC64;

Query Match 68.4%; Score 39; DB 5; Length 72;
Best Local Similarity 66.7%; Pred. No. 6.2;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTTY 9
Db 36 EGDSEIKTHY 44

RESULT 11
Q9ND11 PRELIMINARY; PRT; 74 AA.
ID Q9ND11
AC Q9ND11;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Variant surface protein (Fragment).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]_TaxID=5833;
RP SEQUENCE FROM N.A.
RX STRAIN=WT19;
RX MEDLINE=20372615; PubMed=10910718;
RA Kirchgatter K., Mosbach R., del Portillo H.A.;
RT "Plasmodium falciparum: DBL-1 var sequence analysis in field isolates
RL from central Brazil.";
RL Exp. Parasitol. 95:154-157(2000).
DR EMBL; AF172803; AAF89793.1; -.
FT NON TER 1 1
FT NON TER 74 74
SQ SEQUENCE 74 AA; 8410 MW; 91E750363D125C40 CRC64;

Query Match 68.4%; Score 39; DB 5; Length 74;
Best Local Similarity 66.7%; Pred. No. 6.4;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 QGDSLKTY 9
Db 40 EGDSEIKTH 48

RESULT 12
Q9U7J9 PRELIMINARY; PRT; 179 AA.
ID AC Q9U7J9
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Erythrocyte membrane protein 1 SD105E (Fragment).
GN VAR.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
SEQUENCE FROM N.A.
RP MEDLINE=99404835; PubMed=10477185;
RA Ward C.P., Clotey G.T., Dorris M., Ji D.D., Arnot D.E.;
RT "Analysis of Plasmodium falciparum PfEMP-1/var genes suggests that
  recombination rearranges constrained sequences.";
RL Mol. Biochem. Parasitol. 102:167-177(1999).
DR EMBL; AF127288; AAD52768.1; -
FT NON_TER 1
FT NON_TER 179
SQ SEQUENCE 179 AA; 20577 MW; F625DAA80FC75086 CRC64;

Query Match 68.4%; Score 39; DB 5; Length 179;
Best Local Similarity 66.7%; Pred. No. 16;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 QGDSLKTY 9
Db 32 EGDSEIKNY 40

RESULT 13
ID O61076 PRELIMINARY; PRT; 267 AA.
AC O61076;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Variant-specific surface protein (fragment).
GN VARPH14.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
SEQUENCE FROM N.A.
RP MEDLINE=20183858; PubMed=10717306;
RA Voss T.S., Thompson J.K., Waterkeyn J., Felger I., Weiss N.,
RA Cowman A.F., Beck H.P.;
RT "Genomic distribution and functional characterisation of two distinct
  and conserved Plasmodium falciparum var gene 5' flanking sequences.";
RL Mol. Biochem. Parasitol. 107:103-115(2000).
DR EMBL; AF050739; AAC05219.1; -
FT NON_TER 267
FT NON_TER 267
SQ SEQUENCE 267 AA; 30629 MW; E5E037CB4E43EC1B CRC64;

Query Match 68.4%; Score 39; DB 5; Length 267;
Best Local Similarity 66.7%; Pred. No. 25;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTY 9
Db 144 EGDSEIKTH 152

RESULT 14

```

```

Q9VU83
ID AC Q9VU83 PRELIMINARY; PRT; 516 AA.
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CG10745 protein.
GN CG32130 OR CG10745 OR CG14108.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., Wang X.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003537; AAF49808.1; -
DR FlyBase; FBgn0052130; CG32130.
DR GO; GO:0016329; F:apoptosis regulator activity; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR GO; GO:0006915; P:apoptosis; IEA.
DR InterPro; IPR003103; BAG.
DR Pfam; PF02179; BAG; 1.
SQ SEQUENCE 516 AA; 55885 MW; 7228E69AD53D37F9 CRC64;

Query Match 68.4%; Score 39; DB 5; Length 516;
Best Local Similarity 70.0%; Pred. No. 51;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 QGDSLKTY 10
Db 225 QGQFKTY 234

```

RESULT 15

Q9VU82 PRELIMINARY; PRT; 542 AA.
 ID Q9VU82
 AC Q9VU82;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE CG10745 protein.
 GN CG32130 OR CG10745 OR CG14108.
 OS Drosophila melanogaster (fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkely;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Frannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gortell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*."
 RL Science 287:2185-2195(2000).
 DR EMBL; AF003537; AAF49807.1; --
 DR FlyBase; FBgn0052130; CG32130.
 DR GO; GO:0016329; F:apoptosis regulator activity; IEA.
 DR GO; GO:0005515; F:protein binding; IEA.
 DR GO; GO:0006915; P:apoptosis; IEA.
 DR InterPro; IPR003103; BAG.
 DR Pfam; PF02179; BAG; 1.
 SQ SEQUENCE 542 AA; 58776 MW; 58C0CE3558333557 CRC64;

Query Match 68.4%; Score 39; DB 5; Length 542;
 Best Local Similarity 70.0%; Pred. No. 54;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 QGSLKTYTA 10
 |||
 251 QGQPKTYTA 260

RESULT 16
 Q9VU81 PRELIMINARY; PRT; 609 AA.
 ID Q9VU81; Q95TF3;
 AC Q9VU81; Q95TF3, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE CG32130 protein (GH02003p).
 GN CG32130 OR CG10745 OR CG14108.
 OS Drosophila melanogaster (fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkely;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Frannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gortell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*."
 RL Science 287:2185-2195(2000).
 DR EMBL; AF003537; AAF49807.1; --
 DR FlyBase; FBgn0052130; CG32130.
 DR GO; GO:0016329; F:apoptosis regulator activity; IEA.
 DR GO; GO:0005515; F:protein binding; IEA.
 DR GO; GO:0006915; P:apoptosis; IEA.
 DR InterPro; IPR003103; BAG.
 DR Pfam; PF02179; BAG; 1.
 SQ SEQUENCE FROM N.A.
 RA Celnik S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
 RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
 RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
 RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
 RA Ferreira S.M., Frise E., Galie R.F., Garg N.S., George R.A.,
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
 RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
 RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
 RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
 RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
 RT "Sequencing of *Drosophila melanogaster* genome."

Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

[3]
 SEQUENCE FROM N.A.
 RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
 RA Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
 RA Tupy J.L., Bergman B., Carlson J.W., Celnikier S.E.,
 RA Krommiller B., Marshall B., Millburn G., Richter J., Harris N.,
 RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
 RA Searle S.M.J., Shu S., Smutniak F., Whitfield E.,
 RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
 RT "Annotation of Drosophila melanogaster genome."
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

[4]
 SEQUENCE FROM N.A.
 RA Adams M.D., Celnikier S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

[5]
 SEQUENCE FROM N.A.
 RA FlyBase;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.

[6]
 SEQUENCE FROM N.A.
 RC STRAIN=Berkely;
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise B., George R.,
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Celnikier S.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; A5003537; AAF49809.2; -;
 DR EMBL; AY059435; AAL13341.1; -;
 DR FlyBase; F8gn0052130; CG32130.
 DR GO; GO:0016329; F:apoptosis regulator activity; IEA.
 DR GO; GO:0005515; F:protein binding; IEA.
 DR GO; GO:0006915; F:apoptosis; IEA.
 DR InterPro; IPR003103; BAG.
 DR Pfam; PF02179; BAG; 1.
 DR SMART; SM00264; BAG; 1.
 SQ SEQUENCE 609 AA; 65965 MW; A035C54287B9C354 CRC64;

Query Match 58.4%; Score 39; DB 5; Length 609;
 Best Local Similarity 70.0%; Pred. No. 61;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 QGDSLKYTYA 10
 |||||
 Db 318 QGQFKTYA 327

RESULT 17
 Q8JK71 PRELIMINARY; PRT; 1086 AA.
 AC Q8JK71;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Delta DNA polymerase (EC 2.7.7.7).
 OS Helicobacter pylori strain 33.
 OC Viruses; dsDNA viruses, no RNA stage; Ascoviridae; Ascovirus.
 OX NCBI_TaxID=113368;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C;
 RA Stasiak K., Renault S., Demattei M.V., Bigot Y., Federici B.A.;
 RT "Evolution of Ascoviruses from Iridoviruses."
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.

CC -1- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE = N DIPHOSPHATE
 + [DNA] (N).
 CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.
 CC EMBL; AJ312696; CAC84471.1; -;
 DR GO; GO:0008408; F:3'-5' exonuclease activity; IEA.
 DR GO; GO:0003889; F:alpha DNA polymerase activity; IEA.
 DR GO; GO:0003890; F:beta DNA polymerase activity; IEA.

DR GO; GO:0003891; F:delta DNA polymerase activity; IEA.
 DR GO; GO:0019986; F:deoxycytidyl transferase activity, template. .; IEA.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0003893; F:epsilon DNA polymerase activity; IEA.
 DR GO; GO:0015999; F:eta DNA polymerase activity; IEA.
 DR GO; GO:0003895; F:gamma DNA-directed DNA polymerase activity; IEA.
 DR GO; GO:0016000; F:iota DNA polymerase activity; IEA.
 DR GO; GO:0016450; F:kappa DNA polymerase activity; IEA.
 DR GO; GO:0016449; F:lambda DNA polymerase activity; IEA.
 DR GO; GO:0016448; F:mu DNA polymerase activity; IEA.
 DR GO; GO:0016451; F:nu DNA polymerase activity; IEA.
 DR GO; GO:0000166; F:nucleotide binding; IEA.
 DR GO; GO:0019984; F:sigma DNA polymerase activity; IEA.
 DR GO; GO:0016452; F:theta DNA polymerase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0003894; F:zeta DNA polymerase activity; IEA.
 DR GO; GO:0006260; P:DNA replication; IEA.
 DR InterPro; IPR006172; DNA_pol_B.
 DR InterPro; IPR006134; DNA_pol_B_dom.
 DR InterPro; IPR006133; DNA_pol_B_exo.
 DR Pfam; PF00136; DNA_pol_B; 2.
 DR Pfam; PF03104; DNA_pol_B_exo; 1.
 DR PRINTS; PR00106; DNAPOLB.
 DR SMART; SM00486; POLBc; 1.
 DR PROSITE; PS00116; DNA_POLYMERASE_B; 1.
 KW DNA replication; DNA-binding; DNA-directed DNA polymerase;
 KW Transferase.
 SQ SEQUENCE 1086 AA; 123702 MW; ACE3B2DF0E7B77FB CRC64;

Query Match 58.4%; Score 39; DB 12; Length 1086;
 Best Local Similarity 77.8%; Pred. No. 1.1e+02;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 DSLKTYAS 11
 |||||
 Db 976 DALKTYVS 984

RESULT 18
 Q25734 PRELIMINARY; PRT; 1729 AA.
 AC Q25734;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE PfEMP1 variant 2 of strain MC (Fragment).
 GN MCVAR-2 PFEMP1.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Malayan Camp;
 RX MEDLINE=95330812; PubMed=7541722;
 RA Baruch D.I., Pasloske B.L., Singh H.B., Bi X., Ma X.C., Feldman M.,
 RA Taraschi T.F., Howard R.J.;
 RT "Cloning the P. falciparum gene encoding PfEMP1, a malarial variant
 antigen and adherence receptor on the surface of parasitized human
 erythrocytes."
 RL Cell 82:77-87(1995).
 DR EMBL; U27339; AAA89134.1; -;
 DR PIR; T18396; T18396.
 DR GO; GO:0005539; F:glycosaminoglycan binding; IEA.
 DR GO; GO:0009405; P:pathogenesis; IEA.
 DR InterPro; IPR004258; PFEMP.
 DR Pfam; PF03011; PFEMP; 2.
 FT NON TER 1729 1729
 SQ SEQUENCE 1729 AA; 195156 MW; 7BFERC2131FFBA11 CRC64;

Query Match 58.4%; Score 39; DB 5; Length 1729;
 Best Local Similarity 66.7%; Pred. No. 1.9e+02;
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 QGDSLKTTY 9
Db 151 EGSISIKTHY 159

RESULT 19
Q8IEV1 PRELIMINARY; PRT; 2162 AA.
AC Q8IEV1
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Erythrocyte membrane protein 1 (PFEMP1).
GN VAR.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RA Harris B., Lennard N., Clark L., Line A., Barron A., Corton C.,
RA Berriman M., Pain A., Hall N., Atkin R., Chillingworth C., Doggett J.,
RA Ormond D., Sanders M., Hayes R., Hall S., Quail M., Barrell B.,
RA Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL844509; CAD52143.1; -.
DR GO; GO:0005539; F:glycosaminoglycan binding; IEA.
DR GO; GO:0009405; P:patogenesis; IEA.
DR InterPro; IPR004258; PFEMP.
DR Pfam; PF03011; PFEMP; 2.
SQ SEQUENCE 2162 AA; 247399 MW; 71D641BD28945A0C CRC64;

Query Match 68.4%; Score 39; DB 5; Length 2162;
Best Local Similarity 66.7%; Pred. No. 2.4e+02;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTTY 9
Db 150 EGSISIKTHY 158

RESULT 20
Q9NFB6 PRELIMINARY; PRT; 2163 AA.
AC Q9NFB6
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Var, MAL1P4.01.
GN VAR.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RA Oliver K., Bowman S., Churcher C., Harris B., Barron D.,
RA Quail M., Rajandream M., Barrell B.;
RA Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL031747; CAB89209.1; -.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0005539; F:glycosaminoglycan binding; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR GO; GO:0009405; P:patogenesis; IEA.
DR InterPro; IPR000581; ILVD_EBD_family.
DR InterPro; IPR004258; PFEMP.
DR Pfam; PF03011; PFEMP; 2.
DR PROSITE; PS00886; ILVD_EBD_1; 1.
SQ SEQUENCE 2163 AA; 245805 MW; F5F7AC68E2ABC09 CRC64;

Query Match 68.4%; Score 39; DB 5; Length 2163;
Best Local Similarity 66.7%; Pred. No. 2.4e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 QGDSLKTTY 9
```

```
Db 168 EGSISIKTHY 176

RESULT 21
Q8I495 PRELIMINARY; PRT; 2207 AA.
AC Q8I495
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Erythrocyte membrane protein 1 (PFEMP1).
GN VAR.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RA Devlin K., Baker S., Davies P., Mungal K., Berriman M., Pain A.,
RA Hall N., Bowman S., Churcher C., Quail M., Barrell B.;
RA Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22255708; PubMed=12368867;
RA Hall N., Pain A., Berriman M., Churcher C., Harris B., Harris D.,
RA Mungall K., Bowman S., Atkin R., Baker S., Barron A., Brooks K.,
RA Buckee C.O., Burrows C., Cherevach I., Chillingworth C.,
RA Chillingworth T., Christodoulou Z., Clark L., Clark R., Corton C.,
RA Cronin A., Davies R., Davis P., Dear P., Dearden F., Doggett J.,
RA Fretwell T., Goble A., Goodhead I., Gwilliam R., Hamlin N., Hance Z.,
RA Harper D., Hauser H., Hornsby T., Holroyd S., Horrocks P.,
RA Humphray S., Jagels K., James K.D., Johnson D., Kerkhoun A.,
RA Knights A., Konfortov B., Kyes S., Larke N., Lawson D., Lennard N.,
RA Line A., Maddison M., McLean J., Mooney P., Moule S., Murphy L.,
RA Oliver K., Ormond D., Price C., Quail M.A., Rabinowitsch E.,
RA Rajandream M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M.,
RA Seeger K., Sharp S., Smith R., Squares R., Squares S., Stevens K.,
RA Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J.,
RA Sulston J.E., Craig A., Newbold C., Barrell B.G.;
RT "Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13."
RL Nature 419:527-531(2002).
DR EMBL; AL929351; CAD51367.1; -.
DR GO; GO:0005539; F:glycosaminoglycan binding; IEA.
DR GO; GO:0009405; P:patogenesis; IEA.
DR InterPro; IPR004258; PFEMP.
DR Pfam; PF03011; PFEMP; 2.
SQ SEQUENCE 2207 AA; 251407 MW; 44CD679B3D2FCE24 CRC64;

Query Match 68.4%; Score 39; DB 5; Length 2207;
Best Local Similarity 66.7%; Pred. No. 2.4e+02;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTTY 9
Db 150 EGSISIKTHY 158

RESULT 22
Q25733 PRELIMINARY; PRT; 2924 AA.
AC Q25733
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE PFEMP1 variant 1 of strain MC.
GN MCVAR-1 PFEMP1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Malayan Camp;
RX MEDLINE=95330812; PubMed=7541722;
```

```
RA Baruch D.I., Pasloske B.L., Singh H.B., Bi X., Ma X.C., Feldman M.,
RA Taraschi T.F., Howard R.J.;
RT "Cloning the P. falciparum gene encoding pFEMP1, a malarial variant
RT antigen and adherence receptor on the surface of parasitized human
RT erythrocytes.";
RL Cell 82:77-87(1995)
DR EMBL; U27338; AAB50251.1; -.
DR PIR; T18378; T18378.
DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0005539; F:glycosaminoglycan binding; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR004258; PFEMP.
DR Pfam; PF03011; PFEMP, 2.
DR PROSITE; PS00190; CYTOCHROME_C; 1.
FT VARIANT 104 104 _S -> T.
FT VARIANT 105 105 E -> K.
FT VARIANT 139 139 D -> G.
FT VARIANT 362 362 K -> E.
FT VARIANT 503 503 K -> N.
FT VARIANT 581 581 S -> T.
FT VARIANT 593 593 M -> V.
FT VARIANT 617 617 I -> K.
FT VARIANT 631 631 V -> G.
FT VARIANT 658 658 D -> E.
FT VARIANT 722 722 K -> N.
SQ SEQUENCE 2924 AA; 335860 MW; 7077553BDAC82B26 CRC64;

Query Match 68.4%; Score 39; DB 5; Length 2924;
Best Local Similarity 66.7%; Pred. No. 3.3e+02;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTY 9
Db 151 EGD5IKTHY 159

RESULT 23
Q8IES7 PRELIMINARY; PRT; 9271 AA.
AC Q8IES7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN MAL13P1.19.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RA Harris B., Lennard N., Clark L., Line A., Barron A., Corton C.,
RA Berriman M., Pain A., Hall N., Atkin R., Chillingworth C., Doggett J.,
RA Ormond D., Sanders M., Hayes R., Hall S., Quail M., Barrell B.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL844509; CAD52172.1; -.
DR GO; GO:0004182; P:carboxypeptidase A activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000834; Peptidase_M14.
DR PROSITE; PS00133; CARBOXYPEPT_ZN_2; 1.
KW Hypothetical protein.
SQ SEQUENCE 9271 AA; 1111803 MW; 95438B89C4EAE0E CRC64;

Query Match 68.4%; Score 39; DB 5; Length 9271;
Best Local Similarity 66.7%; Pred. No. 1.1e+03;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 QGDSLKTY 9
Db 915 QGDDIKTYF 923
```

```
RESULT 24
Q9U7K6 PRELIMINARY; PRT; 54 AA.
ID Q9U7K6;
AC Q9U7K6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Erythrocyte membrane protein 1 SD102G (Fragment).
GN VAR.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99404835; PubMed=10477185;
RA Ward C.P., Clotey G.T., Dorris M., Ji D.D., Arnot D.E.;
RT "Analysis of Plasmodium falciparum pFEMP-1/var genes suggests that
RT recombination rearranges constrained sequences.";
RL Mol. Biochem. Parasitol. 102:167-177(1999).
DR EMBL; AF127281; RAD52761.1; -.
FT NON_TER 1 1
FT NON_TER 54 54
SQ SEQUENCE 54 AA; 5986 MW; 43552A60661F4D71 CRC64;

Query Match 66.7%; Score 38; DB 5; Length 54;
Best Local Similarity 60.0%; Pred. No. 7.1;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 QGDSLKTY 10
Db 28 EGD5IRGY 37.

RESULT 25
Q9VJT8 PRELIMINARY; PRT; 56 AA.
ID Q9VJT8;
AC Q9VJT8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE CG15292 protein.
GN CG15292.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Calniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadiou E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Deicher A., Deng C., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.R., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostler C., Gabrielian A.E., Garg N.S., Galbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hosain D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
```

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merklov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapietson M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
DR EMBL; AF003642; AAF53377.1; -;
DR FlyBase; FBgn0040977; CG15292.
SQ SEQUENCE 56 AA; 5926 MW; ECFE0402469F031C CRC64;

Query Match 66.7%; Score 38; DB 5; Length 56;
Best Local Similarity 63.6%; Pred. No. 7.4;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
Db |||:|
14 QGDEMETYEAS 24

Search completed: September 24, 2004, 02:18:24
Job time : 61.717 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2005, 17:57:45 ; Search time 71.1562 Seconds
(without alignments)
79.162 Million cell updates/sec

Title: US-10-088-639A-2_COPY_23_33
Perfect score: 57
Sequence: 1 QGDSLKTYVAS 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	54	94.7	233	2 Q6GMW4	Q6gmw4 homo sapien
2	50	87.7	107	2 Q9NSD6	Q9nsd6 homo sapien
3	46	80.7	81	2 Q72ZE8	Q72ze8 homo sapien
4	44	77.2	161	2 Q96405	Q96405 plasmodium
5	44	77.2	192	2 Q8WPY5	Q8wpys plasmodium
6	44	77.2	2135	2 Q61077	Q61077 plasmodium
7	40	70.2	55	2 Q9TVZ4	Q9tvz4 plasmodium
8	39	68.4	54	2 Q9U7K5	Q9u7k5 plasmodium
9	39	68.4	54	2 Q9U7L1	Q9u7l1 plasmodium
10	39	68.4	60	2 Q9BUJ1	Q9buji plasmodium
11	39	68.4	72	2 Q9ND13	Q9nd13 plasmodium
12	39	68.4	74	2 Q9ND11	Q9nd11 plasmodium
13	39	68.4	179	2 Q9U7J9	Q9u7j9 plasmodium
14	39	68.4	185	2 Q6S911	Q6s911 plasmodium
15	39	68.4	267	2 Q61076	Q61076 plasmodium
16	39	68.4	516	2 Q9VU83	Q9vu83 drosophila
17	39	68.4	542	2 Q9VU82	Q9vu82 drosophila
18	39	68.4	609	2 Q9VU81	Q9vu81 drosophila
19	39	68.4	1040	2 Q704X8	Q704x8 spodoptera
20	39	68.4	1040	2 Q704X9	Q704x9 heliopsis vi
21	39	68.4	1086	2 Q8JK71	Q8jk71 heliopsis vi
22	39	68.4	1729	2 Q25734	Q25734 plasmodium
23	39	68.4	2162	2 Q8IEV1	Q8iev1 plasmodium
24	39	68.4	2163	2 Q9NFB6	Q9nfb6 plasmodium
25	39	68.4	2207	2 Q8I495	Q8i495 plasmodium
26	39	68.4	2924	2 Q25733	Q25733 plasmodium
27	39	68.4	9271	2 Q8IES7	Q8ies7 plasmodium
28	38	66.7	54	2 Q9U7K6	Q9u7k6 plasmodium
29	38	66.7	56	2 Q9VJ78	Q9vj78 drosophila
30	38	66.7	90	2 Q8NKZ5	Q8nkz5 xanthomonas
31	38	66.7	110	2 Q21984	Q21984 streptococ

32	38	66.7	327	2 Q17032	Q17032 anopheles g
33	38	66.7	347	2 Q8I2D6	Q8i2d6 plasmodium
34	38	66.7	402	2 Q7PV90	Q7pv90 anopheles g
35	38	66.7	429	2 Q25470	Q25470 helicobacte
36	38	66.7	429	2 Q9ZL62	Q9zl62 helicobacte
37	38	66.7	624	1 NPR2_YEAST	P39323 saccharomyc
38	38	66.7	624	2 Q6B264	Q6b264 saccharomyc
39	38	66.7	667	2 Q62ES3	Q62es3 burkholderi
40	38	66.7	667	2 Q63YA8	Q63ya8 burkholderi
41	38	66.7	992	2 Q8ILF1	Q8ilf1 plasmodium
42	38	66.7	1697	2 Q8JJG7	Q8jjg7 norwalk-lik
43	37	64.9	228	2 Q7WU96	Q7wu96 cycloclasti
44	37	64.9	297	1 FBX2_MOUSE	Q80uw2 mus musculu
45	37	64.9	313	2 Q8R7K6	Q8r7k6 thermoanaer

ALIGNMENTS

RESULT 1					
Q6GMW4					
ID	Q6GMW4	PRELIMINARY;	PRT;	233 AA.	
AC	Q6GMW4;				
DT	05-JUL-2004 (Tremblrel. 27, Created)				
DT	05-JUL-2004 (Tremblrel. 27, Last sequence update)				
DT	05-JUL-2004 (Tremblrel. 27, Last annotation update)				
DE	Hypothetical protein.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Primary B-Cells;				
RX	MEDLINE=2238857; PubMed=12477932; DOI=10.1073/pnas.242603899;				
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,				
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,				
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,				
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,				
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,				
RA	Stapleton M., Soares M.B., Donald M.P., Casavant T.L., Scheetz T.E.,				
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,				
RA	Raha S.S., Loguillano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,				
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,				
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,				
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,				
RA	Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,				
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,				
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,				
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,				
RA	Krzywicki M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,				
RA	Jones S.J., Marra M.A.;				
RT	"Generation and initial analysis of more than 15,000 full-length human				
RT	and mouse cDNA sequences.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Primary B-Cells;				
RA	Strausberg R.;				
RL	Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; BC073786; AAH73786.1; -.				
DR	GO; GO:0005489; F:electron transporter activity; IEA.				
DR	GO; GO:0006118; P:electron transport; IEA.				
DR	InterPro; IPR003599; IG.				
DR	InterPro; IPR007110; IG-like.				
DR	InterPro; IPR003597; IG.c1.				
DR	InterPro; IPR003006; IG.MHC.				
DR	InterPro; IPR003596; IG.v.				
DR	InterPro; IPR006662; Thired.				
DR	Fam; PF07654; Cl-set; 1.				
DR	Fam; PF00047; ig; 2.				
DR	PRINTS; PR00421; THIOREDOXIN.				
DR	SMART; SM00409; IG; 2.				

```
DR SMART; SM00406; IGL1; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS0835; IG_LIKE; 2.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
KW Hypothetical protein_1.
SQ SEQUENCE 233 AA; 24855 MW; 46267783B9FDE5BD CRC64;

Query Match      94.7%; Score 54; DB 2; Length 233;
Best Local Similarity 90.9%; Pred. No. 0.026;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
Db 42 QGDSLRTYYAS 52

RESULT 2
Q9NSD6 PRELIMINARY; PRT; 107 AA.
AC Q9NSD6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Homo sapiens This CDS feature is included to show the translation of
DE the corresponding V-region. Presently translation qualifiers on
DE V-region features are illegal. (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymphocyte;
RA Hohmann A.;
RL EMBL; L43092; AAA9746.2; -.
DR FIR; S70444; S70444.
DR HSSP; P01709; 2MCG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003596; IG_v.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS0835; IG_LIKE; 1.
FT NON_TER 1
FT NON_TER 107 107
SQ SEQUENCE 107 AA; 11306 MW; A2B04B37187A5F00 CRC64;

Query Match      87.7%; Score 50; DB 2; Length 107;
Best Local Similarity 81.8%; Pred. No. 0.071;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
Db 21 QGDSLRSYYAS 31

RESULT 3
Q7Z2E8 PRELIMINARY; PRT; 81 AA.
AC Q7Z2E8;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Rearranged V131 segment (Rearranged V131 gene segment)
DE (Fragment).
GN Name=VL3L;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Hodgkin lymphoma, and Mantle cell lymphoma;
RA Tinguely M., Rosenquist R., Sundstroem C., Amini R.M., Kuppers R.,

RA Hansmann M.L., Brauning A.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ564423; CAD92030.1; -.
DR EMBL; AJ564424; CAD92031.1; -.
DR HSSP; P01709; 1DCL.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003596; IG_v.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS0835; IG_LIKE; 1.
FT NON_TER 1
FT NON_TER 81 81
SQ SEQUENCE 81 AA; 8702 MW; CFF1D4666B794C9F CRC64;

Query Match      80.7%; Score 46; DB 2; Length 81;
Best Local Similarity 72.7%; Pred. No. 0.33;
Matches 8; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYVAS 11
Db 4 QGDSLRSFYAS 14

RESULT 4
O96405 PRELIMINARY; PRT; 161 AA.
ID O96405
AC O96405;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE Erythrocyte membrane protein 1 (Fragment).
GN Name=FCR3S1.2var9;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RA Chen Q., Wahlgren M.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF039281; AAD02164.1; -.
FT NON_TER 1
FT NON_TER 161 161
SQ SEQUENCE 161 AA; 19005 MW; 164F7D8D9E1AE384 CRC64;

Query Match      77.2%; Score 44; DB 2; Length 161;
Best Local Similarity 77.8%; Pred. No. 1.7;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTYV 9
Db 44 EGD5IKTYV 52

RESULT 5
Q8WPY5 PRELIMINARY; PRT; 192 AA.
ID Q8WPY5
AC Q8WPY5;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2030713; PubMed=12034453; DOI=10.1016/S0166-6851(02)00038-5;
RA Fernandez V., Chen Q., Sundstroem A., Scherf A., Hagblom P.,
RA Wahlgren M.;
RT "Mosaic-like transcription of var genes in single Plasmodium
RL Mol. Biochem. Parasitol. 121:195-203(2002).
DR EMBL; AJ429510; CAD22362.1; -.
KW Hypothetical protein.
```

```
FT NON_TER 1
FT NON_TER 192
SQ SEQUENCE 192 AA; 22196 MW; 022557CE5559DC4C CRC64;

Query Match
Best Local Similarity 77.2%; Score 44; DB 2; Length 192;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTTY 9
Db 43 EGSISIKTTY 51

RESULT 6
O61077 PRELIMINARY; PRT; 2135 AA.
AC O61077;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Variant-specific surface protein (Fragment).
GN Name=varphi17;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20183958; PubMed=10717306; DOI=10.1016/S0166-6851(00)00176-6;
RA Voss T.S., Thompson J.K., Waterkeyn J., Felger I., Weiss N.,
RA Cowman A.F., Beck H.P.;
RT "Genomic distribution and functional characterisation of two distinct
RT and conserved Plasmodium falciparum var gene 5' flanking sequences.";
RL Mol. Biochem. Parasitol. 107:103-115(2000).
DR EMBL; AF050740; AAC05220.1; -.
DR FIR; TI4602; TI4602.
DR GO; GO:0005539; F:glycosaminoglycan binding; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR004258; PFEMP.
DR Pfam; PF03011; PFEMP; 2.
FT NON_TER 2135
SQ SEQUENCE 2135 AA; 242712 MW; C42CEB0C32F4A36F CRC64;

Query Match
Best Local Similarity 77.2%; Score 44; DB 2; Length 2135;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTTY 9
Db 160 EGSISIKTTY 168

RESULT 7
O9TV24 PRELIMINARY; PRT; 55 AA.
AC O9TV24;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Erythrocyte membrane protein 1 SD126F (Erythrocyte membrane protein 1
DE SD102E) (Fragment).
GN Name=var;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99404835; PubMed=10477185; DOI=10.1016/S0166-6851(99)00106-1;
RA Ward C.P., Clotney G.T., Dorris M., Ji D.D., Arnot D.E.;
RT "Analysis of Plasmodium falciparum pFEMP-1/var genes suggests that
RT recombination rearranges constrained sequences.";
RL Mol. Biochem. Parasitol. 102:167-177(1999).
DR EMBL; AF127315; AAD52795.1; -.
DR EMBL; AF127278; AAD52758.1; -.

FT NON_TER 1
FT NON_TER 192
SQ SEQUENCE 192 AA; 22196 MW; 022557CE5559DC4C CRC64;

Query Match
Best Local Similarity 70.2%; Score 40; DB 2; Length 55;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTTY 9
Db 31 EGSISIKTTY 39

RESULT 8
O9U7K5 PRELIMINARY; PRT; 54 AA.
AC O9U7K5;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Erythrocyte membrane protein 1 SD102H (Fragment).
GN Name=var;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99404835; PubMed=10477185; DOI=10.1016/S0166-6851(99)00106-1;
RA Ward C.P., Clotney G.T., Dorris M., Ji D.D., Arnot D.E.;
RT "Analysis of Plasmodium falciparum pFEMP-1/var genes suggests that
RT recombination rearranges constrained sequences.";
RL Mol. Biochem. Parasitol. 102:167-177(1999).
DR EMBL; AF127282; AAD52762.1; -.
FT NON_TER 1
FT NON_TER 54
SQ SEQUENCE 54 AA; 5978 MW; 422FFD9C1FAADA81 CRC64;

Query Match
Best Local Similarity 68.4%; Score 39; DB 2; Length 54;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTTY 9
Db 28 EGSISIKTHY 36

RESULT 9
O9U7L1 PRELIMINARY; PRT; 54 AA.
AC O9U7L1;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Erythrocyte membrane protein 1 SD102B (Fragment).
GN Name=var;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99404835; PubMed=10477185; DOI=10.1016/S0166-6851(99)00106-1;
RA Ward C.P., Clotney G.T., Dorris M., Ji D.D., Arnot D.E.;
RT "Analysis of Plasmodium falciparum pFEMP-1/var genes suggests that
RT recombination rearranges constrained sequences.";
RL Mol. Biochem. Parasitol. 102:167-177(1999).
DR EMBL; AF127275; AAD52755.1; -.
FT NON_TER 1
FT NON_TER 54
SQ SEQUENCE 54 AA; 6050 MW; 94F71D9C1FAAD823 CRC64;

Query Match
Best Local Similarity 68.4%; Score 39; DB 2; Length 54;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
```

```

QY      1 QGDSLKTTY 9
Db      28 EGDSTKTHY 36

RESULT 10
Q9BJJ1 ID Q9BJJ1 PRELIMINARY; PRT; 60 AA.
AC Q9BJJ1
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Variant surface protein (Fragment).
GN Name=var;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RA Nogueira P.A., Wunderlich G., Tada M.S., Costa Jd.N., Menezes M.J.,
RA Scherf A., Pereira-da-Silva L.H.;
RT "Plasmodium falciparum: analysis of transcribed var gene sequences in
RT natural isolates from the Brazilian Amazon region.";
RL Exp. Parasitol. 101:111-120(2002).
DR EMBL; AF333287; AAK19558.1; -.
FT NON_TER 1
FT NON_TER 60
SQ SEQUENCE 60 AA; 7010 MW; FFI73327A53AC20A CRC64;

Query Match 68.4%; Score 39; DB 2; Length 60;
Best Local Similarity 66.7%; Pred. No. 6;
Matches 6; Conservative 2; Mismatches 0; Indels 1; Gaps 0;

QY      1 QGDSLKTTY 9
Db      34 EGDSTKTHY 42

RESULT 11
Q9ND13 ID Q9ND13 PRELIMINARY; PRT; 72 AA.
AC Q9ND13
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Variant surface protein (Fragment).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=20372615; PubMed=10910718; DOI=10.1006/expr.2000.4520;
RA Kirchgatter K., Mosbach R., del Portillo H.A.;
RT "Plasmodium falciparum: DBL-1 var sequence analysis in field isolates
RT from central Brazil.";
RL Exp. Parasitol. 95:154-157(2000).
DR EMBL; AF172801; AAF89791.1; -.
FT NON_TER 1
FT NON_TER 72
SQ SEQUENCE 72 AA; 8198 MW; 715D9AD0DE2D5785 CRC64;

Query Match 68.4%; Score 39; DB 2; Length 72;
Best Local Similarity 66.7%; Pred. No. 7.3;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      1 QGDSLKTTY 9
Db      36 EGDSTKTHY 44

RESULT 12
Q9ND11 ID Q9ND11 PRELIMINARY; PRT; 74 AA.
AC Q9ND11
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Variant surface protein (Fragment).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=20372615; PubMed=10910718; DOI=10.1006/expr.2000.4520;
RA Kirchgatter K., Mosbach R., del Portillo H.A.;
RT "Plasmodium falciparum: DBL-1 var sequence analysis in field isolates
RT from central Brazil.";
RL Exp. Parasitol. 95:154-157(2000).
DR EMBL; AF172803; AAF89793.1; -.
FT NON_TER 1
FT NON_TER 74
SQ SEQUENCE 74 AA; 8410 MW; 91E750363D125C40 CRC64;

Query Match 68.4%; Score 39; DB 2; Length 74;
Best Local Similarity 66.7%; Pred. No. 7.6;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      1 QGDSLKTTY 9
Db      40 EGDSTKTHY 48

RESULT 13
Q9U7J9 ID Q9U7J9 PRELIMINARY; PRT; 179 AA.
AC Q9U7J9
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Erythrocyte membrane protein 1, SDI05E (Fragment).
GN Name=var;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=99404835; PubMed=10477185; DOI=10.1016/S0166-6851(99)00106-1;
RA Ward C.P., Clotney G.T., Dorris M., Ji D.D., Arnot D.E.;
RT "Analysis of Plasmodium falciparum PfEMP-1/var genes suggests that
RT recombination rearranges constrained sequences.";
RL Mol. Biochem. Parasitol. 102:167-177(1999).
DR EMBL; AF127288; AAD52768.1; -.
FT NON_TER 1
FT NON_TER 179
SQ SEQUENCE 179 AA; 20577 MW; F625DAA80FC75086 CRC64;

Query Match 68.4%; Score 39; DB 2; Length 179;
Best Local Similarity 66.7%; Pred. No. 19;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 QGDSLKTTY 9
Db      32 EGDSTKTHY 40

RESULT 14
Q6S911 ID Q6S911 PRELIMINARY; PRT; 185 AA.
AC Q6S911
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Erythrocyte membrane protein 1 (Fragment).
GN Name=var;
OS Plasmodium falciparum.

```

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15122533;
RA Kaestli M., Cortes A., Lagog M., Ott M., Beck H.-P.;
RT "Longitudinal Assessment of Plasmodium falciparum var Gene
Transcription in Naturally Infected Asymptomatic Children in Papua New
Guinea.";
RL J. Infect. Dis. 189:1942-1951(2004).
DR EMBL; AY462689; AAR31929.1; -.
FT InterPro; IPR009085; Hde.
FT NON_TER 1
FT NON_TER 185
SQ SEQUENCE 185 AA; 20466 MW; 5780014F2372C6C4 CRC64;

Query Match 68.4%; Score 39; DB 2; Length 185;
Best Local Similarity 66.7%; Pred. No. 20;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTTY 9
Db 151 EGDSEIKTHY 159

RESULT 15
O61076 PRELIMINARY; PRT; 267 AA.
AC O61076;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Variant-specific surface protein (Fragment).
GN Name=varph14;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20183858; PubMed=10717306; DOI=10.1016/S0166-6851(00)00176-6;
RA Voss T.S., Thompson J.K., Waterkeyn J., Felger I., Weiss N.,
RA Cowman A.F., Beck H.P.;
RT "Genomic distribution and functional characterisation of two distinct
and conserved Plasmodium falciparum var gene 5' flanking sequences.";
RL Mol. Biochem. Parasitol. 107:103-115(2000).
DR EMBL; AF050739; AAC05219.1; -.
FT NON_TER 267
SQ SEQUENCE 267 AA; 30629 MW; E5E037CE4E43EC1B CRC64;

Query Match 68.4%; Score 39; DB 2; Length 267;
Best Local Similarity 66.7%; Pred. No. 30;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGDSLKTTY 9
Db 144 EGDSEIKTHY 152

Search completed: April 28, 2005, 18:24:15
Job time : 73.1562 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 01:32:26 ; Search time 45.0377 Seconds
(without alignments)
43.915 Million cell updates/sec

Title: US-10-088-639a-2_COPY_49_55

Perfect score: 41
Sequence: 1 GNNYRPS 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database : A_Geneseq_29Jan04:.*
1: Geneseq1980s:.*
2: Geneseq1990s:.*
3: Geneseq2000s:.*
4: Geneseq2001s:.*
5: Geneseq2002s:.*
6: Geneseq2003as:.*
7: Geneseq2003bs:.*
8: Geneseq2004s:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41	100.0	249	4 AAB68087	Aab68087 An anti-a
2	37	90.2	182	4 ABB64583	Abb64583 Drosophil
3	37	90.2	182	7 ADD15285	Add15285 Fruitfly
4	37	90.2	865	4 ABB63795	Abb63795 Drosophil
5	37	90.2	930	4 ABUS3195	Abus3195 Human nuc
6	36	87.8	7	3 AAB40011	Aab40011 Anti-hili
7	36	87.8	201	6 AAUI6608	Aui6608 Human nov
8	36	87.8	201	6 ABUS5677	Abus5677 Human nov
9	36	87.8	438	4 ABG10678	Abg10678 Human hum
10	36	87.8	585	2 AAY06297	Aay06297 Human tra
11	36	87.8	588	5 ABG92868	Abg92868 Clock pro
12	36	87.8	591	4 AAUI1686	Aui1686 Human nov
13	36	87.8	591	6 ABUS5255	Abus5255 Human nov
14	36	87.8	599	5 ABG92867	Abg92867 Clock pro
15	36	87.8	602	3 AAY79162	Aay79162 Aryl hydr
16	36	87.8	602	6 ABUS6486	Abus6486 lung canc
17	36	87.8	622	5 ABG92866	Abg92866 Clock pro
18	36	87.8	636	5 ABG92879	Abg92879 BMAL2 (br
19	36	87.8	636	5 ABG92865	Abg92865 Clock pro
20	35	85.4	251	5 ABP45544	Abp45544 Human Bly
21	35	85.4	1182	3 AAB18288	Aab18288 Plasmodiu
22	34	82.9	919	5 ABP35687	Abp35687 Fungal ZB
23	33	80.5	110	4 AAG80220	Aag80220 Human aut
24	33	80.5	155	4 ABG21683	Abg21683 Novel hum
25	33	80.5	157	4 ABG15752	Abg15752 Novel hum

Abg21677 Novel hum
Abg24566 Novel hum
Abg21701 Novel hum
Abg24565 Novel hum
Abg10676 Novel hum
Abg04022 Novel hum
Abg24588 Novel hum
Aao31145 Human CM0
Abg21692 Novel hum
Abp45987 Human Bly
Abp45982 Human Bly
Abp45989 Human Bly
Abp45965 Human Bly
Abg21672 Novel hum
Abp45947 Human Bly
Abg16267 Novel hum
Abg21710 Novel hum
Abp46072 Human Bly
Abg21690 Novel hum
Abg04031 Novel hum
Abg21681 Novel hum
Abg10672 Novel hum
Abg21150 Novel hum
Abg21700 Novel hum
Abg04037 Novel hum
Abg21697 Novel hum
Abg26638 Novel hum
Abg21882 Novel hum
Abg21693 Novel hum
Abg24583 Novel hum
Abg30331 Novel hum
Abg21698 Novel hum
Abg10297 Novel hum
Abg13265 Novel hum
Abg21703 Novel hum
Abg21701 Novel hum
Abg27025 Novel hum
Abg04026 Novel hum
Abg21706 Novel hum
Abb71870 Drosophil
Abg13458 Novel hum
Abg21691 Novel hum
Abg04291 Novel hum
Abg21713 Novel hum
Abg24528 Novel hum
Abg21673 Novel hum
Abg24586 Novel hum
Abg13474 Novel hum
Abg24555 Novel hum
Abg24572 Novel hum
Abg05748 Novel hum
Abg24574 Novel hum
Abg12849 Novel hum
Abg24611 Novel hum
Abg26662 Novel hum
Abg21720 Novel hum
Abg13370 Novel hum
Abg21705 Novel hum
Abg28960 Novel hum
Abg04018 Novel hum
Abg24601 Novel hum
Abg24607 Novel hum
Abg13494 Novel hum
Abg07396 Novel hum
Abg04025 Novel hum
Abg04040 Novel hum
Abg21671 Novel hum
Abg21712 Novel hum
Abg21717 Novel hum
Abg21719 Novel hum
Abg01335 Novel hum
Abg05677 Novel hum
Abg00391 Novel hum

99	33	80.5	529	4	ABG04030	Novel hum
100	33	80.5	529	4	ABG21716	Novel hum
101	33	80.5	536	4	ABG07398	Novel hum
102	33	80.5	536	4	ABG26969	Novel hum
103	33	80.5	538	4	ABG21699	Novel hum
104	33	80.5	538	4	ABG04019	Novel hum
105	33	80.5	547	4	ABG05213	Novel hum
106	33	80.5	550	4	ABG05230	Novel hum
107	33	80.5	550	4	ABG29796	Novel hum
108	33	80.5	550	4	ABG05247	Novel hum
109	33	80.5	555	4	ABG05238	Novel hum
110	33	80.5	561	4	ABG04039	Novel hum
111	33	80.5	564	4	ABG29815	Novel hum
112	33	80.5	567	4	ABG04305	Novel hum
113	33	80.5	571	4	ABG21709	Novel hum
114	33	80.5	572	4	ABG24521	Novel hum
115	33	80.5	574	4	ABG07165	Novel hum
116	33	80.5	575	4	ABG24604	Novel hum
117	33	80.5	580	4	ABG25757	Novel hum
118	33	80.5	580	4	ABG07115	Novel hum
119	33	80.5	580	4	ABG24597	Novel hum
120	33	80.5	580	7	ABG09035	Novel pro
121	33	80.5	584	4	ABG05212	Novel hum
122	33	80.5	588	4	ABG04043	Novel hum
123	33	80.5	590	4	ABG24567	Novel hum
124	33	80.5	594	4	ABG18639	Novel hum
125	33	80.5	594	4	ABG21682	Novel hum
126	33	80.5	595	4	ABG21824	Novel hum
127	33	80.5	595	4	ABG15755	Novel hum
128	33	80.5	595	4	ABG21874	Novel hum
129	33	80.5	595	4	ABG29785	Novel hum
130	33	80.5	595	4	ABG05214	Novel hum
131	33	80.5	598	4	ABG21849	Novel hum
132	33	80.5	603	4	ABG21702	Novel hum
133	33	80.5	603	4	ABG26957	Novel hum
134	33	80.5	611	4	ABG04023	Novel hum
135	33	80.5	615	4	ABG13482	Novel hum
136	33	80.5	621	4	ABG07173	Novel hum
137	33	80.5	627	4	ABG26953	Novel hum
138	33	80.5	628	4	ABG07450	Novel hum
139	33	80.5	629	4	ABG13446	Novel hum
140	33	80.5	630	4	ABG04044	Novel hum
141	33	80.5	631	4	ABG21704	Novel hum
142	33	80.5	631	4	ABG21715	Novel hum
143	33	80.5	633	4	ABG13465	Novel hum
144	33	80.5	633	4	ABG04038	Novel hum
145	33	80.5	638	4	ABG07166	Novel hum
146	33	80.5	643	4	ABG21687	Novel hum
147	33	80.5	643	4	ABG24564	Novel hum
148	33	80.5	643	4	ABG13451	Novel hum
149	33	80.5	652	4	ABG25758	Novel hum
150	33	80.5	652	4	ABG04021	Novel hum

ALIGNMENTS

RESULT 1
AAB68087
ID AAB68087 standard; protein; 249 AA.

XX AAB68087;

XX 09-JUL-2001 (first entry)

DE An anti-alpha6beta4 integrin light chain linked to a heavy chain.

XX Gastrointestinal epithelial tumour cell; alpha6beta4 integrin;

KW tumour-associated antigen; metastatic disease; malignant disease;

XX tumour typing; tumour screening; tumour.

OS Synthetic.

OS Macaca fascicularis.

XX	Key	Location/Qualifiers
FH	Region	23. .33
FT		/note= "Complementarity determining region (CDR) 1 of the light chain"
FT	Region	49. .55
FT		/note= "Complementarity determining region (CDR) 2 of the light chain"
FT	Region	88. .98
FT		/note= "Complementarity determining region (CDR) 3 of the light chain"
FT	Peptide	110. .127
FT		/note= "linker"
FT	Region	158. .162
FT		/note= "Complementarity determining region (CDR) 1 of the heavy chain"
FT	Region	177. .193
FT		/note= "Complementarity determining region (CDR) 1 of the heavy chain"
FT	Region	226. .238
FT		/note= "Complementarity determining region (CDR) 1 of the heavy chain"

WO200130854-A2.

03-MAY-2001.

26-OCT-2000; 2000WO-SE002082.

28-OCT-1999; 99SE-00003895.

(ACTI-) ACTIVE BIOTECH AB.

Brodin TN, Karlstroem PJ, Ohlsson IG, Tordsson MJ, Kearney PP;

Nilson BHK;

WPI; 2001-308619/32.

N-PSDB; AAF84797.

Novel antibody for diagnosis, treatment of human metastatic and malignant diseases, has binding structure for target structure displayed on cell surface of human gastrointestinal epithelial tumor and normal cells.

Claim 1; Page 55-56; 75pp; English.

The present sequence represents a Monkey antibody light chain linked to a heavy chain. The antibody binds to a target structure displayed in and on the cell surface of human gastrointestinal epithelial tumour cells and in a subpopulation of normal human gastrointestinal epithelial cells. The target structure especially comprises alpha6beta4 integrin. This is a tumour-associated antigen. The antibody, and its fragments, are useful for treating conditions based on anti-angiogenic mechanism in humans. They are useful for treating human metastatic and malignant disease, for in vitro, in vivo diagnosis and prognosis of human malignant disease, comprising tumour typing, tumour screening, tumour diagnosis and prognosis and monitoring premalignant conditions. Quantitative in vivo diagnosis is carried out by determining the localization of antibody to tumour deposits in humans

Sequence 249 AA;

Query Match 100.0%; Score 41; DB 4; Length 249;

Best Local Similarity 100.0%; Pred. No. 15;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GNNYRPS 7

Db 49 GNNYRPS 55

RESULT 2

ABB64583

ID ABB64583 standard; protein; 182 AA.

XX ABB64583;
XX 26-MAR-2002 (first entry)
XX Drosophila melanogaster polypeptide SEQ ID NO 20541.
XX DE Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical.
XX KW Drosophila melanogaster.
XX OS WO200171042-A2.
XX PN 27-SEP-2001.
XX PD 23-MAR-2001; 2001WO-US009231.
XX PF 23-MAR-2000; 2000US-0191637P.
XX PR 11-JUL-2000; 2000US-00614150.
XX PA (PEKE) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX DR WPI; 2001-656860/75.
XX DR N-PSDB; ABL08686.
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX PT genes from Drosophila and for elucidating cell signaling and cell-cell
XX PT interactions.
XX PS Disclosure; SEQ ID NO 20541; 21pp + Sequence Listing; English.
XX CC The invention relates to an isolated nucleic acid detection reagent
XX CC capable of detecting 1000 or more genes from Drosophila. The invention is
XX CC useful in developmental biology and in elucidating cell signalling and
XX CC cell-cell interactions in higher eukaryotes for the development of
XX CC insecticides, therapeutics and pharmaceutical drugs. The invention
XX CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX CC sequences (ABL01840-ABL6175) and the encoded proteins (ABB57737-
XX CC ABB72072). The sequence data for this patent did not form part of the
XX CC printed specification, but was obtained in electronic format directly
XX CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX CC Sequence 182 AA;
Query Match 90.2%; Score 37; DB 4; Length 182;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GNNYRP 6
DB 134 GNNYRP 139
RESULT 3
ADD15285
ID ADD15285 standard; protein; 182 AA.
AC ADD15285;
XX 15-JAN-2004 (first entry)
XX DT Fruitfly odourant receptor protein (SeqId 37).
XX DE odourant receptor; fruitfly; mating; repulsion; flight; insect damage;
XX KW odourant receptor; fruitfly; mating; repulsion; flight; insect damage;
XX KW disease spread; pesticide; insect management program.
XX XX Drosophila melanogaster.
XX OS WO2003020913-A2.
XX PN Venter JC, Adams M, Li PWD, Myers EW;
PI

PD 13-MAR-2003.
XX 04-SEP-2002; 2002WO-US028315.
XX PR 04-SEP-2001; 2001US-0317401P.
XX PA (SENT-) SENTIGEN CORP.
XX PI Lee KJ, Ong J, Nguyen TT, Kloss B;
XX WIPI; 2003-300885/29.
XX DR Novel Anopheles gambiae odourant receptor polypeptides and nucleic acid
XX PT encoding the polypeptides, useful as targets for identifying pest control
XX PT agents.
XX PS Claim 1; SEQ ID NO 37; 172pp; English.
XX CC This invention relates to novel Anopheles gambiae odourant receptor genes
XX CC and encoded proteins thereof. Specifically, it refers to the isolated
XX CC genes of the African malaria mosquito that are related to the 'classical'
XX CC Drosophila odourant receptor genes, and compounds that bind to and
XX CC modulate these receptors cause various behavioural responses such as
XX CC mating, repulsion or flight. Accordingly, the present invention describes
XX CC such compounds (both natural and synthetic) that are useful for
XX CC attracting insects to traps or to localised toxins, for repelling insects
XX CC from individuals or populated residential areas, or for interfering with
XX CC the function of olfactory system such that insects are unable to locate
XX CC food and hosts. As such, these compounds can be used to control insect
XX CC damage and the spread of disease, and will significantly reduce
XX CC dependence on toxic pesticides having a direct and immediate impact on
XX CC coordinated insect management programs. This polypeptide sequence is a
XX CC Drosophila melanogaster odourant receptor protein, which is homologous to
XX CC the African malaria mosquito proteins of the invention.
XX CC Sequence 182 AA;
Query Match 90.2%; Score 37; DB 7; Length 182;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GNNYRP 6
DB 134 GNNYRP 139
RESULT 4
ABB63795
ID ABB63795 standard; protein; 865 AA.
XX ABB63795;
XX AC ABB63795;
XX DT 26-MAR-2002 (first entry)
XX DE Drosophila melanogaster polypeptide SEQ ID NO 18177.
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical.
XX OS Drosophila melanogaster.
XX PN WO200171042-A2.
XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-US009231.
XX PR 23-MAR-2000; 2000US-0191637P.
XX PR 11-JUL-2000; 2000US-00614150.
XX PA (PEKE) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
PI

XX WPI; 2001-656860/75.
DR N-PSDB; ABL07898.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
PS Disclosure; SEQ ID NO 18177; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 865 AA;
Query Match 90.2%; Score 37; DB 4; Length 865;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 GNNYRP 6
Db 424 GNNYRP 429
|||||
XX
RESULT 5
ABUS3195
ID ABUS3195 standard; protein; 930 AA.
XX
AC ABUS3195;
XX
DT 14-APR-2003 (first entry)
XX
DE Human nucleic acid management-associated DKFZphtes3_2m18 homologue.
XX
KW Human; gene therapy; vaccine; disease treatment; detection.
XX
OS Homo sapiens.
XX
XX WO200112659-A2.
XX
XX 22-FEB-2001.
XX
XX 18-AUG-2000; 2000WO-IB001496.
XX
XX 18-AUG-1999; 99US-0149499P.
XX
XX 28-SEP-1999; 99US-0156503P.
XX
XX (GEHU-) GERMAN HUMAN GENOME PROJECT.
XX
XX Wiemann S;
XX
XX WPI; 2001-327840/34.
XX
PT Nucleic acids having the sequences of clones isolated from libraries of
PT different human tissues, useful in recombinant DNA methodologies.
XX
XX Example III; Page 809; 1095pp; English.
XX
CC This invention describes novel polynucleotides and polypeptides isolated
CC from human cDNA libraries which can be used for gene therapy or in
CC vaccines. The polynucleotides of the invention and antibodies encoded by
CC them may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate polypeptide expression. The products of the
CC invention may also be used to identify modulators of expression and
CC activity and to down regulate expression and activity. The antibodies of

CC the invention may also be used as diagnostic agents for detecting the
CC presence of polypeptides in samples. This sequence represents a homologue
CC of a polypeptide described in the disclosure of the invention
XX
SQ Sequence 930 AA;
Query Match 90.2%; Score 37; DB 4; Length 930;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 GNNYRP 6
Db 843 GNNYRP 848
|||||
XX
RESULT 6
AAB40011
ID AAB40011 standard; peptide; 7 AA.
XX
AC AAB40011;
XX
DT 05-FEB-2001 (first entry)
XX
DE Anti-hIL12 antibody light chain CDR2 amino acid sequence SEQ ID 527.
XX
KW Human; neutralising antibody; interleukin-12; IL-12; antiinflammatory;
KW complementarity determining region; CDR; antirheumatic; antiarthritic;
KW antisclerotic; neuroprotective; antipsoriatic; antiasthmatic; cardiant;
KW antiparasitic; antibacterial; immunosuppressive; Crohn's disease;
KW multiple sclerosis; rheumatoid arthritis.
XX
OS Homo sapiens.
XX
XX WO2000056772-A1.
XX
XX 28-SEP-2000.
XX
XX 24-MAR-2000; 2000WO-US007946.
XX
XX 25-MAR-1999; 99US-0126603P.
XX
XX (BADI) BASF AG.
XX (GEMY) GENETICS INST INC.
XX
XX Salfeld JG, Roguska M, Paskind M, Banerjee S, Tracey DE, White M;
XX Kaymakalan Z, Labkovsky B, Sakorafas P, Friedrich S, Myles A;
XX Veldman GM, Venturini A, Warne NW, Widom A, Elvin JG, Duncan AR;
XX Derbyshire EJ, Carmen S, Smith S, Holtet TL, Du Fou SL;
XX
XX WPI; 2000-638250/61.
XX
XX New human antibody specific for human interleukin-12 (IL-12) used to
XX treat disorders characterized by aberrant IL-12 expression e.g. Crohn's
XX disease and multiple sclerosis.
XX
XX Claim 33; Fig 2G; 377pp; English.
XX
XX This invention relates to a new human antibody specific for human
XX interleukin-12 (IL-12). The invention also includes antigen binding
XX portions that bind to IL-12. Sequences AAB39485-B39516 represent human
XX anti-IL-12 antibody heavy and light chain complementarity determining
XX region (CDR) amino acid sequences, and also includes variable region
XX amino acid sequences. Other variable region amino acid sequences are
XX given in AAB39517-B39560 and AAB40068-B40149. Sequences AAB39561-B39771
XX represent anti-IL-12 CDR3 related amino acid sequences. AAB39772-B40063
XX represent other CDR sequences. Light chain CDR3 consensus sequences are
XX given in AAB40064-B40067. Primers used in the identification and
XX construction of the antibodies of the invention are given in AAC61062-
XX C61071. The antibody of the invention is a neutralising antibody and has
XX antirheumatic; antirheumatic; antisclerotic; antiparasitic;
XX neuroprotective; antipsoriatic; antiasthmatic; cardiant; antiparasitic;
XX antibacterial and immunosuppressive activity. The antibodies or antigen-
XX binding fragments are useful in the treatment of disorders associated

CC with detrimental release of human IL-12, especially Crohn's disease,
CC multiple sclerosis and rheumatoid arthritis. They can also be used in the
CC manufacture of a pharmaceutical composition to treat human IL-12
CC disorders

SQ Sequence 7 AA;

Query Match 87.8%; Score 36; DB 3; Length 7;
Best Local Similarity 85.7%; Pred. NO. 1.4e+06;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
||:||||
DB 1 GNDYRPS 7

RESULT 7

AAU16608

ID AAU16608 standard; protein; 201 AA.

XX

AC AAU16608;

XX 07-NOV-2001 (first entry)

DT Human novel secreted protein, Seq ID 1561.

DE

KW Human; immunosuppressive; antiarthritic; antirheumatic; cytostatic;
KW cardiac; vasotropic; cerebroprotective; nootropic; neuroprotective;
KW antibacterial; virucide; fungicide; ophthalmological; vulnerary;
KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW cerebral ischaemia; angiogenesis; nervous system disorder;
KW Alzheimer's disease; infection; ocular disorder; corneal infection;
KW wound healing; epithelial cell proliferation; skin ageing; food additive;
KW preservative; antiproliferative.

XX

OS Homo sapiens.

XX

PN WO200155322-A2.

XX

PD 02-AUG-2001.

XX

PF 17-JAN-2001; 2001WO-US001341.

XX

PR 31-JAN-2000; 2000US-0179065P.

PR

PR 04-FEB-2000; 2000US-0180628P.

PR

PR 24-FEB-2000; 2000US-0184664P.

PR

PR 02-MAR-2000; 2000US-0186350P.

PR

PR 16-MAR-2000; 2000US-0189874P.

PR

PR 17-MAR-2000; 2000US-0190076P.

PR

PR 18-APR-2000; 2000US-0198123P.

PR

PR 19-MAY-2000; 2000US-0205515P.

PR

PR 07-JUN-2000; 2000US-0209467P.

PR

PR 28-JUN-2000; 2000US-0214886P.

PR

PR 30-JUN-2000; 2000US-0215135P.

PR

PR 07-JUL-2000; 2000US-0216647P.

PR

PR 17-JUL-2000; 2000US-0216880P.

PR

PR 11-JUL-2000; 2000US-0217487P.

PR

PR 11-JUL-2000; 2000US-0217496P.

PR

PR 14-JUL-2000; 2000US-0218290P.

PR

PR 26-JUL-2000; 2000US-0220963P.

PR

PR 26-JUL-2000; 2000US-0220964P.

PR

PR 14-AUG-2000; 2000US-0224518P.

PR

PR 14-AUG-2000; 2000US-0224519P.

PR

PR 14-AUG-2000; 2000US-0225213P.

PR

PR 14-AUG-2000; 2000US-0225214P.

PR

PR 14-AUG-2000; 2000US-0225266P.

PR

PR 14-AUG-2000; 2000US-0225267P.

PR

PR 14-AUG-2000; 2000US-0225268P.

PR

PR 14-AUG-2000; 2000US-0225270P.

PR

PR 14-AUG-2000; 2000US-0225447P.

PR

PR 14-AUG-2000; 2000US-0225757P.

PR

PR 14-AUG-2000; 2000US-0225758P.

PR

PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241828P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.

PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239935P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 XX
 PA (ROSE/) ROSEN C A.
 PA (RUBE/) RUBEN S M.
 PA (BARA/) BARASH S C.
 XX
 PI Rosen CA, Ruben SM, Barash SC;
 XX
 DR WPI; 2003-147444/14.
 DR N-PSDB; ABX73936.
 XX
 XX New polypeptides and nucleic acids, useful in gene therapy for treating,
 PT inhibiting or preventing e.g. neural, immune system, muscular,
 PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or
 PT renal disorders.
 XX
 PS Claim 11; SEQ ID NO 1561; 402pp; English.
 XX
 CC The invention relates to human novel polypeptides and their associated
 CC polynucleotides. The polypeptides and polynucleotides are useful in gene
 CC therapy for treating, inhibiting or preventing neural disorders, immune
 CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis
 CC and multiple sclerosis), muscular disorders, respiratory diseases (e.g.
 CC nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,
 CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
 CC (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left
 CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage
 CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and
 CC leukaemia), inflammatory diseases (e.g. septic shock, bursitis and
 CC appendicitis), allergic reactions and conditions (e.g. asthma), blood
 CC related disorders (e.g. thrombosis, atherosclerosis and myocardial
 CC infarction) and cancerous diseases. Sequences ABU54914-ABU55699 and
 CC ABU55748 represent human novel polypeptides of the invention
 XX
 SQ Sequence 201 AA;
 Query Match 87.8%; Score 36; DB 6; Length 201;
 Best Local Similarity 85.7%; Pred. NO. 1e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GNNYRPS 7
 Db 124 GSNYRPS 130
 RESULT 9
 ABG10678
 ID ABG10678 standard; protein; 438 AA.
 XX
 AC ABG10678;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #10669.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US008631.
 XX
 PR 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI; 2001-639362/73.
 DR N-PSDB; AAS74865.
 XX
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 PS Claim 20; SEQ ID NO 41037; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 438 AA;
 Query Match 87.8%; Score 36; DB 4; Length 438;
 Best Local Similarity 85.7%; Pred. NO. 2.2e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GNNYRPS 7
 Db 174 GNDYRPS 180
 RESULT 10
 AAY06297
 ID AAY06297 standard; protein; 585 AA.
 XX
 AC AAY06297;
 XX
 DT 23-AUG-1999 (first entry)
 XX
 DE Human transcription regulator MOP9.
 XX
 KW MOP9; member of the PAS superfamily; bHLH-PAS; human;
 KW transcription regulator; hypoxia.
 XX
 OS Homo sapiens.
 XX
 PN WO928464-A2.
 XX
 PD 10-JUN-1999.
 XX

PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US008631.
 XX
 PR 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI; 2001-639362/73.
 DR N-PSDB; AAS74865.
 XX
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 PS Claim 20; SEQ ID NO 41037; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 438 AA;
 Query Match 87.8%; Score 36; DB 4; Length 438;
 Best Local Similarity 85.7%; Pred. NO. 2.2e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GNNYRPS 7
 Db 174 GNDYRPS 180
 RESULT 10
 AAY06297
 ID AAY06297 standard; protein; 585 AA.
 XX
 AC AAY06297;
 XX
 DT 23-AUG-1999 (first entry)
 XX
 DE Human transcription regulator MOP9.
 XX
 KW MOP9; member of the PAS superfamily; bHLH-PAS; human;
 KW transcription regulator; hypoxia.
 XX
 OS Homo sapiens.
 XX
 PN WO928464-A2.
 XX
 PD 10-JUN-1999.
 XX

PF 27-NOV-1998; 98WO-US025314.
XX
PR 28-NOV-1997; 97US-0066863P.
XX
PA (WISC) WISCONSIN ALUMNI RES FOUND.
XX
PI Bradfield CA, Gu YZ, Hogenesch JB;
XX
XX WPI; 1999-371120/31.
DR N-PSDB; AAX58988.
XX
XX Developmental signal transduction associated proteins.
PT
PS Claim 5; Page 106; 106pp; English.
XX
CC The present sequence represents MOP9, a novel member of the PAS
CC superfamily, where PAS stands for PER/ARNT/SIM domains. MOP9 cDNA (see
CC AAX58988) was cloned from human brain mRNA. MOP9 displays extended
CC homology to MOP3 (see AAY06291). It pairs with CLOCK and MOP4 (see
CC AAY06292) and binds an E-box element with flanking region specificity.
CC The invention provides novel MOP 2-9 nucleic acids (see AAX58981-88) and
CC proteins (see AAY06289-97). These are useful in a variety of research,
CC diagnostic and therapeutic applications. Several of the MOPs are alpha-
CC class hypoxia-inducible factors. Others are involved in circadian signal
CC transduction
XX
XX
SQ Sequence 585 AA;
Query Match 87.8%; Score 36; DB 2; Length 585;
Best Local Similarity 85.7%; Pred. No. 2.9e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GNNYRPS 7
Db 118 GSNYRPS 124
RESULT 11
ID ABG92868 standard; protein; 588 AA.
XX
AC ABG92868;
XX
DT 29-AUG-2003 (revised)
DT 19-NOV-2002 (first entry)
XX
DE Clock protein BAML2 (brain-muscle-Arnt-like protein 2) #4.
XX
KW Human; clock protein BMAL2; brain-muscle-Arnt-like protein 2; insomnia;
KW sleeping disorder; non-24-hour sleep; sleep-phase forward;
KW retreat syndrome; time-zone variation syndrome.
XX
OS Homo sapiens.
XX
PN WO200264785-A1.
XX
PD 22-AUG-2002.
XX
PF 23-AUG-2001; 2001WO-JP007197.
XX
PR 13-FEB-2001; 2001JP-00035743.
XX
PA (NISC-) JAPAN.SCI & TECHNOLOGY CORP.
XX
PI Fukuda Y, Okano T;
PI
DR WPI; 2002-667007/71.
DR N-PSDB; ABS68504.
XX
XX Clock gene Baml2 and expressed clock protein BMAL2 important in clock
PT oscillation mechanism and relating to circadian rhythm, used in diagnosis
PT of and developing drugs for insomnia and other sleeping disorders.
XX

PS Claim 1; Page 102-105; 187pp; Japanese.
XX
CC The invention relates to a DNA sequence encoding clock protein BMAL2
CC (brain-muscle-Arnt-like protein 2). The gene and protein are applicable
CC in diagnosis of and development of drugs for insomnia and other sleeping
CC disorders e.g. non-24-hour sleep, sleep-phase forward or retreat syndrome
CC and time-zone variation syndrome. ABG92865-ABG92879 represent BMAL2 amino
CC acid sequences of the invention. (Updated on 29-AUG-2003 to standardise
CC OS field)
XX
XX Sequence 588 AA;
Query Match 87.8%; Score 36; DB 5; Length 588;
Best Local Similarity 85.7%; Pred. No. 2.9e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GNNYRPS 7
Db 121 GSNYRPS 127
RESULT 12
AAU16186
ID AAU16186 standard; protein; 591 AA.
XX
AC AAU16186;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human novel secreted protein, Seq ID 1139.
XX
KW Human; immunosuppressive; antiarthritic; antirheumatic; cytostatic;
KW cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective;
KW antibacterial; virucide; fungicide; ophthalmological; vulnerary;
KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW cerebral ischaemia; angiogenesis; nervous system disorder;
KW Alzheimer's disease; infection; ocular disorder; corneal infection;
KW wound healing; epithelial cell proliferation; skin ageing; food additive;
KW preservative; antiproliferative.
XX
OS Homo sapiens.
XX
PN WO200155322-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001341.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0218880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225474P.
PR 14-AUG-2000; 2000US-0225575P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234597P.
PR 25-SEP-2000; 2000US-0234598P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.

PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-02559678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
PA Rosen CA, Barash SC, Ruben SM;
PI
XX WPI: 2001-488783/53.
XX N-PSDB; AAS26173.
DR
XX
PT New nucleic acid molecules encoding 461 human secreted proteins for
PT diagnosing, preventing, treating or ameliorating medical conditions and
XX used as food additives or preservatives.
PS Claim 11; SEQ ID NO 1139; 980pp; English.
XX
CC The invention relates to isolated nucleic acid molecules and their
CC encoded secreted proteins. The nucleic acids and proteins are used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
CC in diagnosing a pathological condition or susceptibility to a
CC pathological condition. Antibodies to the proteins can also be used in
CC alleviating symptoms associated with the disorders and in diagnostic
CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbant assays
CC (ELISA). Disorders which are diagnosed or treated include autoimmune
CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,
CC nervous system disorders e.g. Alzheimer's disease, infections caused by
CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection,
CC and many other disorders listed in the specification. The polypeptides
CC can also be used to aid wound healing and epithelial cell proliferation,
CC to prevent skin aging due to sunburn, to maintain organs before
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The polypeptides can also be used
CC as a food additive or preservative to increase or decrease storage

CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
CC minerals, cofactors and other nutritional components. The present
CC sequence represents a novel secreted protein of the invention. Note: The
CC sequence data for this patent did not form part of the printed

Query Match 87.8%; Score 36; DB 4; Length 591;
Best Local Similarity 85.7%; Pred. No. 2.9e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
Db 124 GSNYRPS 130

RESULT 13
ABU55255
ID ABU55255 standard; protein; 591 AA.

XX AC ABU55255;

DT 18-MAR-2003 (first entry)

XX DE Human novel polypeptide #342.

XX KW Human; neural disorder; immune system disorder; renal disorder;
KW muscular disorder; respiratory disease; reproductive disorder;
KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
KW hyperproliferative disorder; inflammatory disease; allergic reaction;
KW blood related disorder; cancer; immunosuppressive; antiinflammatory;
KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;
KW haemostatic; antiarteriosclerotic.

OS Homo sapiens.

XX PN US2002132753-A1.

XX PD 19-SEP-2002.

XX PF 17-JAN-2001; 2001US-00764864.

XX PR 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 28-JUN-2000; 2000US-0214886P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225757P.

PR 14-AUG-2000; 2000US-0225758P.

PR 22-AUG-2000; 2000US-0226868P.

PR 30-AUG-2000; 2000US-0228924P.

PR 01-SEP-2000; 2000US-0229287P.

PR 01-SEP-2000; 2000US-0229343P.

PR 29-SEP-2000; 2000US-0236368P.

PR 29-SEP-2000; 2000US-0236369P.

PR 29-SEP-2000; 2000US-0236370P.

PR 02-OCT-2000; 2000US-0236802P.

PR 02-OCT-2000; 2000US-0237037P.

PR 02-OCT-2000; 2000US-0237038P.

PR 02-OCT-2000; 2000US-0237039P.

PR 02-OCT-2000; 2000US-0237040P.

PR 13-OCT-2000; 2000US-0239935P.

PR 20-OCT-2000; 2000US-0240960P.

PR 20-OCT-2000; 2000US-0241785P.

PR 20-OCT-2000; 2000US-0241809P.

PR 01-NOV-2000; 2000US-0244617P.

PR 17-NOV-2000; 2000US-0249299P.

PR 08-DEC-2000; 2000US-0251856P.

PR 08-DEC-2000; 2000US-0251868P.

PR 08-DEC-2000; 2000US-0251869P.

XX (ROSE/) ROSEN C A.

PA (RUBE/) RUBEN S M.

PA (BARA/) BARASH S C.

XX PI Rosen CA, Ruben SM, Barash SC;

XX WPI; 2003-147444/14.

DR N-PSDB; ABX73514.

XX New polypeptides and nucleic acids, useful in gene therapy for treating,

PT inhibiting or preventing e.g. neural, immune system, muscular,

PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or

PT renal disorders.

XX Claim 11; SEQ ID NO 1139; 402pp; English.

XX The invention relates to human novel polypeptides and their associated

CC polynucleotides. The polypeptides and polynucleotides are useful in gene

CC therapy for treating, inhibiting or preventing neural disorders, immune

CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis

CC and multiple sclerosis), muscular disorders, respiratory diseases (e.g.

CC nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,

CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders

CC (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left

CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage

CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and

CC leukaemia), inflammatory diseases (e.g. septic shock, bursitis and

CC appendicitis), allergic reactions and conditions (e.g. asthma), blood

CC related disorders (e.g. thrombosis, atherosclerosis and myocardial

CC infarction) and cancerous diseases. Sequences ABU54914-ABU55699 and

CC ABU55748 represent human novel polypeptides of the invention

XX Sequence 591 AA;

QY

Db

RESULT 14

ABG92867

ID ABG92867 standard; protein; 599 AA.

XX AC ABG92867;

XX DT 29-AUG-2003 (revised)

DT 19-NOV-2002 (first entry)

XX DE Clock protein BAML2 (brain-muscle-Arnt-like protein 2) #3.

XX DE Human; clock protein BAML2; brain-muscle-Arnt-like protein 2; insomnia;

XX KW

Query Match 87.8%; Score 36; DB 6; Length 591;

Best Local Similarity 85.7%; Pred. No. 2.9e+02;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNYRPS 7

Db 124 GSNYRPS 130

KW sleeping disorder; non-24-hour sleep; sleep-phase forward;
 KW retreat syndrome; time-zone variation syndrome.
 XX Homo sapiens.
 XX WO200264785-A1.
 XX PD 22-AUG-2002.
 XX PF 23-AUG-2001; 2001WO-JP007197.
 XX PR 13-FEB-2001; 2001JP-00035743.
 XX PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.
 XX PI Fukada Y, Okano T;
 XX WPI; 2002-667007/71.
 XX N-PSDB; AB68503.
 XX Clock gene Bmal2 and expressed clock protein BMAL2 important in clock
 PT oscillation mechanism and relating to circadian rhythm, used in diagnosis
 PT of and developing drugs for insomnia and other sleeping disorders.
 XX Claim 1; Page 94-97; 187pp; Japanese.
 XX The invention relates to a DNA sequence encoding clock protein BMAL2
 CC (brain-muscle-Arnt-like protein 2). The gene and protein are applicable
 CC in diagnosis of and development of drugs for insomnia and other sleeping
 CC disorders e.g. non-24-hour sleep, sleep-phase forward or retreat syndrome
 CC and time-zone variation syndrome. ABG92865-ABG92879 represent BMAL2 amino
 CC acid sequences of the invention. (Updated on 29-AUG-2003 to standardise
 CC OS field)
 XX Sequence 599 AA;

Query Match 87.8%; Score 36; DB 5; Length 599;
 Best Local Similarity 85.7%; Pred. No. 2.9e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GNNYRPS 7
 DB 132 GSNYRPS 138
 RESULT 15
 AAY79162
 ID AAY79162 standard; protein; 602 AA.
 XX AAY79162;
 XX 05-JUN-2000 (first entry)
 DT Aryl hydrocarbon receptor nuclear translocator-4 (ARNT4).
 DE Aryl hydrocarbon receptor nuclear translocator-4; ARNT4; human;
 KW angio genesis; antiarteriosclerotic; antitumour; atherosclerosis; tumour;
 KW vascular disease; vulnerability; cardiac; vasotropic; cerebroprotective;
 KW gene therapy; circadian rhythm.
 XX Homo sapiens.
 XX WO200009657-A2.
 XX PD 24-FEB-2000.
 XX PF 13-AUG-1999; 99WO-US018539.
 XX PR 14-AUG-1998; 98US-0096515P.
 XX (HARD) HARVARD COLLEGE.
 XX Lee M, Maemura K, Hiesh C;

XX WPI; 2000-205996/18.
 DR N-PSDB; AA294062.
 XX Modulation of angiogenesis in mammals, useful for treating e.g.
 PT atherosclerosis, tumors, wounds, vascular disease, hypoxic tissue damage,
 PT ischemia, balloon angioplasty, frostbite, gangrene or poor circulation.
 XX Claim 26; Page 31-32; 57pp; English.
 XX The present sequence is that of human aryl hydrocarbon receptor nuclear
 CC translocator-4 (ARNT4), a novel basic helix-loop-helix (bHLH)/PAS protein
 CC that is an important regulator of vascular endothelial cell growth factor
 CC (VEGF) gene expression, especially in the vascular system. ARNT4
 CC interacts with endothelial PAS domain protein 1 (EPAS1, see MAY79161),
 CC forming a heterodimer that binds to the hypoxia responsive element of the
 CC VEGF gene. A claimed method of inhibiting angiogenesis in a mammal
 CC comprises administering to the mammal a compound which inhibits binding
 CC of EPAS1 to ARNT4. The compound may comprise amino acid residues 75-128,
 CC 155-207 or 232-3847 of ARNT4 or is the full-length protein. Modulation of
 CC ARNT4 production or activity can also be used to regulate circadian
 CC rhythms, e.g. by forming a heterodimer with Clock, or to treat circadian
 CC rhythm disorders. Expression of ARNT4 through the methods of gene therapy
 CC using ARNT4 DNA can be used to promote new blood vessel formation, e.g.
 CC to promote angiogenesis in wound healing and organ transplantation, to
 CC treat peripheral vascular disease, cerebral vascular disease, hypoxic
 CC tissue damage, coronary vascular disease, or following transient
 CC ischaemic attacks, vascular graft surgery, balloon angioplasty,
 CC frostbite, gangrene or poor circulation
 XX Sequence 602 AA;
 Query Match 87.8%; Score 36; DB 3; Length 602;
 Best Local Similarity 85.7%; Pred. No. 3e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GNNYRPS 7
 DB 135 GSNYRPS 141
 RESULT 16
 ABUS6486
 ID ABUS6486 standard; protein; 602 AA.
 XX ABUS6486;
 XX 02-APR-2003 (first entry)
 DT Lung cancer-associated polypeptide #79.
 DE Lung cancer-associated polypeptide; cytostatic; emphysema;
 KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
 KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
 KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
 KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
 XX Unidentified.
 XX WO200286443-A2.
 XX 31-OCT-2002.
 XX 18-APR-2002; 2002WO-US012476.
 XX 18-APR-2001; 2001US-0284770P.
 XX 10-MAY-2001; 2001US-0290492P.
 XX 09-NOV-2001; 2001US-0339245P.
 XX 13-NOV-2001; 2001US-0350666P.
 XX 29-NOV-2001; 2001US-0334370P.
 XX 12-APR-2002; 2002US-0372246P.
 XX (EOSB-) EOS BIOTECHNOLOGY INC.

XX PI Aziz N, Murray R;
XX DR WPI; 2003-093161/08.
XX DR N-PSDB; ABX76210.
XX PT Detecting a lung cancer-associated transcript in a cell from a patient
XX PT for treating lung cancer, by contacting a biological sample from the
XX PT patient with a polynucleotide that exhibits increased or decreased
XX PT expression in lung cancer.
XX PS Claim 27; Page 250; 453pp; English.
XX CC The invention relates to a method for detecting a lung cancer-associated
XX CC transcript in a cell from a patient, comprising contacting a biological
XX CC sample from the patient with a polynucleotide that selectively hybridizes
XX CC to a sequence that is at least 80 % identical to a gene that exhibits
XX CC increased or decreased expression in lung cancer samples. Lung cancer-
XX CC associated polynucleotides and polypeptides are used for identifying a
XX CC compound that modulates a lung cancer-associated polypeptide, for
XX CC inhibiting proliferation of a lung cancer-associated cell to treat lung
XX CC cancer in a patient and for treating a mammal having lung cancer by
XX CC administering a modulatory compound identified. The methods are useful
XX CC for treating lung cancer, such as small cell lung cancer, non-small cell
XX CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
XX CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
XX CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
XX CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
XX CC for diagnostic purposes and as targets for screening for therapeutic
XX CC compounds that modulate lung cancer, such as antibodies. Sequences
XX CC ABU56408-ABU56745 represent lung cancer-associated polypeptides of the
XX CC invention
XX SQ Sequence 602 AA;
Query Match 87.8%; Score 36; DB 6; Length 602;
Best Local Similarity 85.7%; Pred. No. 3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GNNYRPS 7
Db 135 GSNYRPS 141
RESULT 17
ABG92866
ID ABG92866 standard; protein; 622 AA.
XX AC ABG92866;
XX DT 29-AUG-2003 (revised)
XX DT 19-NOV-2002 (first entry)
XX DE Clock protein BAML2 (brain-muscle-Arnt-like protein 2) #2.
XX DE Human; clock protein BAML2; brain-muscle-Arnt-like protein 2; insomnia;
XX KW sleeping disorder; non-24-hour sleep; sleep-phase forward;
XX KW retreat syndrome; time-zone variation syndrome.
XX OS Homo sapiens.
XX XX WO200264785-A1.
XX PD 22-AUG-2002.
XX PF 23-AUG-2001; 2001WO-JP007197.
XX XX 13-FEB-2001; 2001JP-00035743.
XX PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX PI Fukada Y, Okano T;
XX PI

DR WPI; 2002-667007/71.
DR N-PSDB; ABS68502.
XX CC Clock gene Baml2 and expressed clock protein BMAL2 important in clock
XX PT oscillation mechanism and relating to circadian rhythm, used in diagnosis
XX PT of and developing drugs for insomnia and other sleeping disorders.
XX PS Claim 1; Page 85-88; 187pp; Japanese.
XX CC The invention relates to a DNA sequence encoding clock protein BMAL2
XX CC (brain-muscle-Arnt-like protein 2). The gene and protein are applicable
XX CC in diagnosis of and development of drugs for insomnia and other sleeping
XX CC disorders e.g. non-24-hour sleep, sleep-phase forward or retreat syndrome
XX CC and time-zone variation syndrome. ABG92865-ABG92879 represent BMAL2 amino
XX CC acid sequences of the invention. (Updated on 29-AUG-2003 to standardise
XX CC OS field)
XX SQ Sequence 622 AA;
Query Match 87.8%; Score 36; DB 5; Length 622;
Best Local Similarity 85.7%; Pred. No. 3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GNNYRPS 7
Db 155 GSNYRPS 161
RESULT 18
ABG92879
ID ABG92879 standard; protein; 636 AA.
XX AC ABG92879;
XX DT 19-NOV-2002. (first entry)
XX DE BMAL2 (brain-muscle-Arnt-like protein 2)-related protein #4.
XX XX Human; clock protein BMAL2; brain-muscle-Arnt-like protein 2; insomnia;
XX KW sleeping disorder; non-24-hour sleep; sleep-phase forward;
XX KW retreat syndrome; time-zone variation syndrome.
XX OS Unidentified.
XX XX WO200264785-A1.
XX PD 22-AUG-2002.
XX PF 23-AUG-2001; 2001WO-JP007197.
XX XX 13-FEB-2001; 2001JP-00035743.
XX PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX PI Fukada Y, Okano T;
XX XX WPI; 2002-667007/71.
XX PT Clock gene Baml2 and expressed clock protein BMAL2 important in clock
XX PT oscillation mechanism and relating to circadian rhythm, used in diagnosis
XX PT of and developing drugs for insomnia and other sleeping disorders.
XX PS Example 1; Fig 3; 187pp; Japanese.
XX CC The invention relates to a DNA sequence encoding clock protein BMAL2
XX CC (brain-muscle-Arnt-like protein 2). The gene and protein are applicable
XX CC in diagnosis of and development of drugs for insomnia and other sleeping
XX CC disorders e.g. non-24-hour sleep, sleep-phase forward or retreat syndrome
XX CC and time-zone variation syndrome. ABG92865-ABG92879 represent BMAL2 amino
XX CC acid sequences of the invention
XX SQ Sequence 636 AA;
XX

Query Match 87.8%; Score 36; DB 5; Length 636;
Best Local Similarity 85.7%; Pred. No. 3.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
Db 169 GSNYRPS 175

RESULT 19

ABG92865
ID ABG92865 standard; protein; 636 AA.

XX AC ABG92865;

DT 29-AUG-2003 (revised)

DT 19-NOV-2002 (first entry)

DE Clock protein BAML2 (brain-muscle-Arnt-like protein 2) #1.

XX Human; clock protein BMAL2; brain-muscle-Arnt-like protein 2; insomnia;
KW sleeping disorder; non-24-hour sleep; sleep-phase forward;
KW retreat syndrome; time-zone variation syndrome.

XX OS Homo sapiens.

PN WO200264785-A1.

PD 22-AUG-2002.

PF 23-AUG-2001; 2001WO-JP007197.

PR 13-FEB-2001; 2001JP-00035743.

PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.

XX PI Fukada Y, Okano T;

XX DR WPI; 2002-667007/71.

DR N-PSDB; ABS68501.

XX Clock gene Baml2 and expressed clock protein BMAL2 important in clock
PT oscillation mechanism and relating to circadian rhythm, used in diagnosis
PT of and developing drugs for insomnia and other sleeping disorders.

XX PS Claim 1; Page 76-80; 187pp; Japanese.

XX CC The invention relates to a DNA sequence encoding clock protein BMAL2
CC (brain-muscle-Arnt-like protein 2). The gene and protein are applicable
CC in diagnosis of and development of drugs for insomnia and other sleeping
CC disorders e.g. non-24-hour sleep, sleep-phase forward or retreat syndrome
CC and time-zone variation syndrome. ABG92865-ABG92879 represent BMAL2 amino
CC acid sequences of the invention. (Updated on 29-AUG-2003 to standardise
CC OS field)

XX SQ Sequence 636 AA;

Query Match 87.8%; Score 36; DB 5; Length 636;
Best Local Similarity 85.7%; Pred. No. 3.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
Db 169 GSNYRPS 175

RESULT 20

ABP45544
ID ABP45544 standard; protein; 251 AA.

XX AC ABP45544;

XX DT 19-AUG-2002 (first entry).

XX

DE Human BlyS binding scFv SEQ ID 1555.

XX

KW BlyS; B lymphocyte stimulator; TNF superfamily; human; cytostatic;
KW tumour necrosis factor; B cell proliferation; B cell differentiation;
KW immunosuppressive; immunostimulant; immunomodulatory; antirheumatic;
KW antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;
KW systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;
KW common variable immunodeficiency; acquired immunodeficiency syndrome.

XX OS Homo sapiens.

XX PN WO200202641-A1.

XX PD 10-JAN-2002.

PF 15-JUN-2001; 2001WO-US019110.

XX PR 16-JUN-2000; 2000US-0212210P.

PR 17-OCT-2000; 2000US-0240816P.

PR 16-MAR-2001; 2001US-0276248P.

PR 21-MAR-2001; 2001US-0277379P.

PR 25-MAY-2001; 2001US-0293499P.

XX

PA (HUMA-) HUMAN GENOME SCI INC.

PA (CAMB-) CAMBRIDGE ANTIBODY TECHNOLOGY.

XX

PI Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;

XX

DR WPI; 2002-114799/15.

XX

PT Antibodies against B Lymphocyte Stimulating polypeptides, useful for the

PT diagnosis and treatment of cancers and immune disorders.

XX

PS Claim 1; Page 2258-2259; 3148pp; English.

XX

XX CC This invention describes novel antibodies that immunospecifically bind to
CC B Lymphocyte Stimulator (BlyS) polypeptides. BlyS is a member of the
CC tumour necrosis factor (TNF) super family and induces B cell
CC proliferation and differentiation. The antibodies of the invention have
CC cytostatic, immunosuppressive, immunostimulant, immunomodulatory,
CC antirheumatic and antiAIDS activity and can be used in vaccines to
CC inhibit the expression and activity of BlyS. The antibodies bind to BlyS
CC and so may be used to detect and quantitate the presence of BlyS in
CC biological samples and may be used in this way to diagnose disease
CC associated with aberrant expression of BlyS. They may also be
CC administered to treat diseases associated with aberrant BlyS expression
CC and activity such as cancer, immune, and autoimmune disorders and
CC diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis,
CC immunodeficiency (e.g. common variable immunodeficiency (CVID) and
CC acquired immunodeficiency syndrome (AIDS)). ABP43990-ABP47228 represent
CC the antibodies and fragments of the antibodies described in the method of
CC the invention

XX SQ Sequence 251 AA;

Query Match 85.4%; Score 35; DB 5; Length 251;
Best Local Similarity 85.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
Db 191 GKNYRPS 197

RESULT 21

AAB18288

ID AAB18288 standard; protein; 1182 AA.

XX

AC AAB18288;

XX

DT 07-NOV-2000 (first entry)

DE Plasmodium falciparum chromosome 2 related protein SEQ ID NO:146.
 XX Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine;
 KW antimalarial; malaria; protozoicide; infection; insecticide.
 XX Plasmodium falciparum.
 OS WO200025728-A2.
 PN 11-MAY-2000.
 XX 05-NOV-1999; 99WO-US026796.
 PF 05-NOV-1998; 98US-0107131P.
 PR (HOFF/) HOFFMAN S.
 XX (CARU/) CARUCCI D.
 PA (GARD/) GARDNER M.
 PA (VENT/) VENTER J C.
 PI Hoffman S, Carucci D, Gardner M, Venter JC;
 XX WPI; 2000-365347/31.
 DR Proteins encoded by chromosome 2 of the human malarial parasite,
 XX Plasmodium falciparum, useful as antimalarial vaccines and in the
 PT diagnosis of P.falciparum infection.
 PT Disclosure; Page 347-350; 577pp; English.
 PS The present invention describes proteins and their fragments (I) encoded
 CC by chromosome 2 of the human malarial parasite, Plasmodium falciparum.
 CC Also described are: (I) nucleotide sequences (II) encoding (I); and (2)
 CC vaccines against P. falciparum infection comprising (I) or (II). (I) and
 CC (II) are useful for the development of vaccines against P. falciparum
 CC infection. (I) and polyclonal antisera or a monoclonal antibody raised to
 CC immunogens comprising the sequences of (I), are useful in the detection
 CC of infection with P. falciparum. Furthermore, (I) (especially when they
 CC are rifins or secreted or membrane proteins) can aid the identification
 CC of drugs to treat or prevent P. falciparum infection, or they can be used
 CC to identify drug resistance in P. falciparum. Sequencing of the
 CC Plasmodium chromosome 2 and the subsequent identification of proteins
 CC encoded by it will help to expand our understanding of parasite biology,
 CC a process hampered by the complexity of the parasitic lifecycle, and
 CC provide new targets for vaccine and drug development. Parasite resistance
 CC to drugs and mosquito resistance to insecticides have led to a resurgence
 CC of malaria in many parts of the world, and there is a pressing need for
 CC vaccines and new drugs. AAA70078 to AAA70287 and AAB18144 to AAB18352
 CC represent nucleotide and protein sequences given in the present
 CC invention, but which are not specifically mentioned within the
 CC specification
 XX
 SQ Sequence 1182 AA;
 Query Match 85.4%; Score 35; DB 3; Length 1182;
 Best Local Similarity 100.0%; Pred. No. 8.6e-02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 NNYRPS 7
 DB 624 NNYRPS 629
 |||||
 RESULT 22
 ID ABP35687
 AC ABP35687 standard; protein; 919 AA.
 XX
 AC ABP35687;
 XX
 DT 24-JUL-2002 (first entry)
 XX
 DE Fungal ZBC protein sequence #113.
 XX

KW Secondary metabolite; fungus; ZBC gene; zinc binuclear cluster protein;
 KW antibacterial; beta-lactam; anti-hypercholesterolaemic; lovastatin;
 KW mevastatin; immunosuppressant; cyclosporin A; ergot alkaloid; ergotamine;
 KW angiogenesis inhibitor; ovalicin; glucan synthase inhibitor; gliotoxin;
 KW fungal toxin; cell surface receptor; plant growth regulator; pigment;
 KW insecticide; antineoplastic.
 XX
 OS Unidentified.
 XX WO200224865-A2.
 PN 28-MAR-2002.
 PD 19-SEP-2001; 2001WO-US029288.
 PF 19-SEP-2000; 2000US-0233564P.
 PR (MICK-) MICROBIA INC.
 XX Holtzman D, Madden K, Maxon M, Sherman A;
 PI WPI; 2002-352005/38.
 DR N-PSDB; ABN79876.
 DR New method for improving the production of a secondary metabolite e.g.
 XX antineoplastic agent, ergot alkaloid from a fungus involves modulation of
 PT the expression of at least one zinc binuclear cluster protein gene.
 PT Disclosure; SEQ ID NO 238; 49pp + Sequence Listing; English.
 PS The invention relates to improving the production of a secondary
 CC metabolite by a fungus. This involves modulating the expression of at
 CC least one ZBC (zinc binuclear cluster protein) gene in a manner to
 CC improve the yield of the secondary metabolite. Methods of the invention
 CC may be used for improving the production of the secondary metabolite e.g.
 CC antibacterial (such as beta-lactam), an anti-hypercholesterolaemic (such
 CC as lovastatin or mevastatin), an immunosuppressant (such as cyclosporin A),
 CC an ergot alkaloid (such as ergotamine), an angiogenesis inhibitor (such
 CC as ovalicin), a glucan synthase inhibitor, gliotoxin family of compounds,
 CC a fungal toxin, a modulator of cell surface receptor signalling, a plant
 CC growth regulator, a pigment, an insecticide, or an antineoplastic
 CC compound. The method results in a decrease in fermentor run-time, a
 CC decrease in the size of the fermentor required for the production of
 CC equivalent amounts of the secondary metabolite, or a decrease in the
 CC biomass required for the production, which translates into decreased
 CC waste that must be handled in downstream processing. The sequences given
 CC in records ABP35575-ABP35722 represent ZBC proteins. Note: The sequence
 CC data for this patent did not form part of the printed specification, but
 CC was obtained directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 919 AA;
 Query Match 82.9%; Score 34; DB 5; Length 919;
 Best Local Similarity 85.7%; Pred. No. 1e+03;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GNNYRPS 7
 DB 130 GNNYRPS 136
 |||||
 RESULT 23
 ID AAG80220
 AC AAG80220 standard; protein; 110 AA.
 XX
 AC AAG80220;
 XX
 DT 22-JAN-2002 (first entry)
 XX
 DE Human autoantibody MiCA-9 variable region heavy chain.
 XX
 KW Autoantigen; fusion protein; islet cell antigen; MiCA autoantibody;

KW Glutamate decarboxylase; diabetes mellitus type I; stiff-man syndrome;
 KW polyglandular autoimmune syndrome; autoimmune disorder; IA2; GAD65;
 KW variable region; heavy chain; MICA-9.
 XX Homo sapiens.
 XX EPI149914-A2.
 XX 31-OCT-2001.
 XX 29-MAR-2001; 2001EP-00107702.
 XX 10-APR-2000; 2000DE-01017782.
 XX 25-MAY-2000; 2000DE-01025840.
 XX (LABO-) LABOR KOCH MERK GMBH.
 XX Richter W, Rickert M, Rapp I, Dangel W;
 XX WPI; 2001-640702/74.
 XX N-PSDB; AAI68769.
 XX New fusion protein, useful for diagnosis of diabetes type I and other
 PT metabolic diseases, is reactive with autoantibodies against both
 PT glutamate decarboxylase and islet cell antigen.
 XX Disclosure; Page 19; 68pp; German.
 XX This invention describes a novel fusion protein (I) that has, at its N-
 CC terminus, one or more epitopes that bind specifically to autoantibodies
 CC (AAB) against the islet cell antigen IA2 and, at its C-terminus, one or
 CC more epitopes that bind specifically to antibodies (Ab) directed against
 CC the glutamate decarboxylase GAD65. (I), also nucleic acid (II) encoding
 CC it, vectors containing (II) and transformed cells, are useful for
 CC diagnosis and prognosis of diabetes mellitus type 1, stiff-man syndrome,
 CC polyglandular autoimmune syndrome or other autoimmune conditions
 CC associated with AAB against GAD65 or IA2. (I) provides a rapid and simple
 CC diagnosis of high specificity and sensitivity, capable of recognizing
 CC antibodies against both IA2 and GAD65, simultaneously. Unlike known
 CC fusions, where the GAD65 component is at the N-terminus, (I) contains
 CC correctly folded conformational epitopes that can react with most MICA
 CC autoantibodies. This sequence represents the human autoantibody MICA-9
 CC variable region heavy chain used in the method of the invention
 XX Sequence 110 AA;
 SQ Query Match 80.5%; Score 33; DB 4; Length 110;
 Best Local Similarity 85.7%; Pred. No. 2e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GNNYRPS 7
 DB 51 GNNLRPS 57
 RESULT 24
 ABG21683
 ID ABG21683 standard; protein; 155 AA.
 XX ABG21683;
 XX 18-FEB-2002 (first entry)
 DE Novel human diagnostic protein #21674.
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX Homo sapiens.
 XX WO200175067-A2.
 XX 11-OCT-2001.
 PD

XX 30-MAR-2001; 2001WO-US008631.
 XX 31-MAR-2000; 2000US-00540217.
 XX 23-AUG-2000; 2000US-00649167.
 XX (HYSE-) HYSEQ INC.
 XX Drmanac RT, Liu C, Tang YT;
 XX WPI; 2001-639362/73.
 XX N-PSDB; AAS85870.
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX Claim 20; SEQ ID NO 52042; 103pp; English.
 XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 155 AA;
 SQ Query Match 80.5%; Score 33; DB 4; Length 155;
 Best Local Similarity 71.4%; Pred. No. 2.8e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GNNYRPS 7
 DB 83 GNDYRPT 89
 RESULT 25
 ABG15752
 ID ABG15752 standard; protein; 157 AA.
 XX ABG15752;
 XX 18-FEB-2002 (first entry)
 DE Novel human diagnostic protein #15743.
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX Homo sapiens.
 XX WO200175067-A2.
 XX 11-OCT-2001.
 XX 30-MAR-2001; 2001WO-US008631.
 PF

XX 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
DR N-PSDB; AAS79939.
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX Claim 20; SEQ ID NO 46111; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 157 AA;

Query Match 80.5%; Score 33; DB 4; Length 157;
Best Local Similarity 71.4%; Pred.No. 2.8e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
Db 120 GNDYRPT 126

Search completed: September 24, 2004, 01:53:06
Job time : 53.0377 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 01:53:42 ; Search time 11.8868 Seconds
(without alignments)
30.402 Million cell updates/sec

Title: US-10-088-639A-2_COPY_49_55
Perfect score: 41
Sequence: 1 GNNYRPS 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : Issued Patents_AA:*

- 1: /cgn2_6/prodata/2/iaa/5A_COMB.pep.*
- 2: /cgn2_6/prodata/2/iaa/5B_COMB.pep.*
- 3: /cgn2_6/prodata/2/iaa/6A_COMB.pep.*
- 4: /cgn2_6/prodata/2/iaa/6B_COMB.pep.*
- 5: /cgn2_6/prodata/2/iaa/PCrUS_COMB.pep.*
- 6: /cgn2_6/prodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	87.8	602	4	US-09-374-454-19
2	33	80.5	157	4	US-09-489-039A-11409
3	33	80.5	324	4	US-09-489-039A-11252
4	32	78.0	132	2	US-08-345-321-4
5	32	78.0	159	4	US-09-621-976-4449
6	32	78.0	223	4	US-09-107-532A-7214
7	32	78.0	348	1	US-08-035-392-2
8	32	78.0	348	1	US-08-504-511A-2
9	32	78.0	430	1	US-08-035-392-4
10	32	78.0	430	1	US-08-504-511A-4
11	32	78.0	442	4	US-09-540-236-3486
12	32	78.0	552	4	US-09-835-654-2
13	32	78.0	720	4	US-09-480-921B-8
14	32	78.0	1115	7	US-09-376-594-893
15	31	75.6	7	2	US-08-672-345C-38
16	31	75.6	7	2	US-08-672-345C-74
17	31	75.6	7	3	US-09-214-095D-38
18	31	75.6	7	3	US-09-214-095D-74
19	31	75.6	74	4	US-09-621-976-4090
20	31	75.6	75	4	US-09-621-976-5643
21	31	75.6	109	2	US-08-672-345C-3
22	31	75.6	109	2	US-08-672-345C-93
23	31	75.6	109	3	US-09-214-095D-3
24	31	75.6	109	3	US-09-214-095D-121
25	31	75.6	330	4	US-09-489-039A-10928
26	31	75.6	388	2	US-08-994-772-2
27	31	75.6	388	2	US-09-207-844-2
28	28	68.3	70	3	US-09-383-212-4
29	28	68.3	70	3	US-09-383-212-4
30	28	68.3	70	3	US-09-383-212-4
31	28	68.3	70	3	US-09-383-212-4
32	28	68.3	70	3	US-09-383-212-4
33	28	68.3	70	3	US-09-383-212-4
34	28	68.3	70	3	US-09-383-212-4
35	28	68.3	70	3	US-09-383-212-4
36	28	68.3	70	3	US-09-383-212-4
37	28	68.3	70	3	US-09-383-212-4
38	28	68.3	70	3	US-09-383-212-4
39	28	68.3	70	3	US-09-383-212-4
40	28	68.3	70	3	US-09-383-212-4
41	28	68.3	70	3	US-09-383-212-4
42	28	68.3	70	3	US-09-383-212-4
43	28	68.3	70	3	US-09-383-212-4
44	28	68.3	70	3	US-09-383-212-4
45	28	68.3	70	3	US-09-383-212-4
46	28	68.3	70	3	US-09-383-212-4
47	28	68.3	70	3	US-09-383-212-4
48	28	68.3	70	3	US-09-383-212-4
49	28	68.3	70	3	US-09-383-212-4
50	28	68.3	70	3	US-09-383-212-4
51	28	68.3	70	3	US-09-383-212-4
52	28	68.3	70	3	US-09-383-212-4
53	28	68.3	70	3	US-09-383-212-4
54	28	68.3	70	3	US-09-383-212-4
55	28	68.3	70	3	US-09-383-212-4
56	28	68.3	70	3	US-09-383-212-4
57	28	68.3	70	3	US-09-383-212-4
58	28	68.3	70	3	US-09-383-212-4
59	28	68.3	70	3	US-09-383-212-4
60	28	68.3	70	3	US-09-383-212-4
61	28	68.3	70	3	US-09-383-212-4
62	28	68.3	70	3	US-09-383-212-4
63	28	68.3	70	3	US-09-383-212-4
64	28	68.3	70	3	US-09-383-212-4
65	28	68.3	70	3	US-09-383-212-4
66	28	68.3	70	3	US-09-383-212-4
67	28	68.3	70	3	US-09-383-212-4
68	28	68.3	70	3	US-09-383-212-4
69	28	68.3	70	3	US-09-383-212-4
70	28	68.3	70	3	US-09-383-212-4
71	28	68.3	70	3	US-09-383-212-4
72	28	68.3	70	3	US-09-383-212-4
73	28	68.3	70	3	US-09-383-212-4
74	28	68.3	70	3	US-09-383-212-4
75	28	68.3	70	3	US-09-383-212-4
76	28	68.3	70	3	US-09-383-212-4
77	28	68.3	70	3	US-09-383-212-4
78	28	68.3	70	3	US-09-383-212-4
79	28	68.3	70	3	US-09-383-212-4
80	28	68.3	70	3	US-09-383-212-4
81	28	68.3	70	3	US-09-383-212-4
82	28	68.3	70	3	US-09-383-212-4
83	28	68.3	70	3	US-09-383-212-4
84	28	68.3	70	3	US-09-383-212-4
85	28	68.3	70	3	US-09-383-212-4
86	28	68.3	70	3	US-09-383-212-4
87	28	68.3	70	3	US-09-383-212-4
88	28	68.3	70	3	US-09-383-212-4
89	28	68.3	70	3	US-09-383-212-4
90	28	68.3	70	3	US-09-383-212-4
91	28	68.3	70	3	US-09-383-212-4
92	28	68.3	70	3	US-09-383-212-4
93	28	68.3	70	3	US-09-383-212-4
94	28	68.3	70	3	US-09-383-212-4
95	28	68.3	70	3	US-09-383-212-4
96	28	68.3	70	3	US-09-383-212-4
97	28	68.3	70	3	US-09-383-212-4
98	28	68.3	70	3	US-09-383-212-4
99	28	68.3	70	3	US-09-383-212-4
100	28	68.3	70	3	US-09-383-212-4

Sequence 5033, Ap
Sequence 4, Appli
Sequence 4, Appli
Sequence 4, Appli
Sequence 4, Appli
Sequence 405, App
Sequence 11, Appl
Sequence 80, Appl
Sequence 80, Appl
Sequence 12288, A
Sequence 108, App
Sequence 108, App
Sequence 2, Appli
Sequence 2592, Ap
Sequence 2, Appli
Sequence 23, Appl
Sequence 26847, A
Sequence 7526, Ap
Sequence 8, Appli
Sequence 10, Appl
Sequence 12, Appl
Sequence 59, Appl
Sequence 8455, Ap
Sequence 4186, A
Sequence 1394, A
Sequence 15, Appl
Sequence 15, Appl
Sequence 127, App
Sequence 24655, A
Sequence 5009, Ap
Sequence 4, Appli
Sequence 4, Appli
Sequence 25610, A
Sequence 20, Appl
Sequence 12, Appl
Sequence 12, Appl
Sequence 11, Appl
Sequence 11, Appl
Sequence 11, Appl
Sequence 5841, Ap
Sequence 14, Appl
Sequence 98, Appl
Sequence 98, Appl
Sequence 98, Appl
Sequence 98, Appl
Sequence 98, Appl
Sequence 98, Appl
Sequence 98, Appl
Sequence 15, Appl
Sequence 15, Appl
Sequence 98, Appl
Sequence 98, Appl
Sequence 6327, Ap
Sequence 22867, A
Sequence 4, Appli
Sequence 4, Appli
Sequence 9337, Ap
Sequence 43, Appl
Sequence 2, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 31, Appl
Sequence 4, Appli
Sequence 4, Appli

```
101 28 68.3 117 4 US-09-543-681A-4837 Sequence 4837, Ap
102 28 68.3 126 3 US-08-793-426A-7 Sequence 7, Appli
103 28 68.3 126 3 US-09-294-565-7 Sequence 7, Appli
104 28 68.3 161 4 US-09-489-847-188 Sequence 188, App
105 28 68.3 214 4 US-09-489-847-336 Sequence 336, App
106 28 68.3 214 4 US-09-252-991A-16879 Sequence 16879, A
107 28 68.3 224 4 US-09-489-039A-13337 Sequence 13337, A
108 28 68.3 229 4 US-09-540-236-3318 Sequence 3318, Ap
109 28 68.3 250 4 US-10-039-785-50 Sequence 50, Appl
110 28 68.3 283 4 US-09-328-352-4670 Sequence 4670, Ap
111 28 68.3 284 1 US-08-390-858B-36 Sequence 36, Appl
112 28 68.3 319 4 US-09-134-001C-3888 Sequence 3888, Ap
113 28 68.3 329 4 US-09-252-991A-17777 Sequence 17777, A
114 28 68.3 331 3 US-08-793-426A-2 Sequence 2, Appli
115 28 68.3 331 3 US-09-294-565-2 Sequence 2, Appli
116 28 68.3 496 4 US-09-252-991A-26506 Sequence 26506, A
117 28 68.3 498 3 US-09-238-303-8 Sequence 8, Appli
118 28 68.3 498 4 US-09-946-239-8 Sequence 8, Appli
119 28 68.3 520 4 US-09-252-991A-30677 Sequence 30677, A
120 28 68.3 523 4 US-09-489-039A-9202 Sequence 9202, Ap
121 28 68.3 609 3 US-08-941-445A-7 Sequence 7, Appli
122 28 68.3 618 4 US-09-934-901-18 Sequence 18, Appl
123 28 68.3 685 3 US-08-947-965-72 Sequence 72, Appl
124 28 68.3 686 3 US-08-947-965-70 Sequence 70, Appl
125 28 68.3 686 3 US-08-947-965-73 Sequence 73, Appl
126 28 68.3 719 5 PCT-US94-01624-31 Sequence 31, Appl
127 28 68.3 731 4 US-09-107-532A-6999 Sequence 6999, Ap
128 28 68.3 731 1 US-08-021-601-4 Sequence 4, Appli
129 28 68.3 735 1 US-08-082-849B-4 Sequence 4, Appli
130 28 68.3 735 5 PCT-US94-01624-4 Sequence 4, Appli
131 28 68.3 735 5 PCT-US94-01624-3 Sequence 3, Appli
132 28 68.3 772 4 US-09-252-991A-30121 Sequence 30121, A
133 28 68.3 813 4 US-09-540-236-3472 Sequence 3472, Ap
134 28 68.3 803 1 US-08-021-601-12 Sequence 12, Appl
135 28 68.3 903 1 US-08-082-849B-12 Sequence 12, Appl
136 28 68.3 903 5 PCT-US94-01624-12 Sequence 12, Appl
137 28 68.3 923 4 US-09-397-885-1 Sequence 1, Appli
138 28 68.3 923 4 US-09-969-362-1 Sequence 1, Appli
139 28 68.3 1436 4 US-10-080-505-13 Sequence 13, Appl
140 28 68.3 1436 4 US-08-264-093-19 Sequence 19, Appl
141 27 65.9 7 4 US-08-918-148-10 Sequence 10, Appl
142 27 65.9 14 1 US-08-179-481-53 Sequence 53, Appl
143 27 65.9 68 4 US-09-328-352-8224 Sequence 8224, Ap
144 27 65.9 75 4 US-09-134-001C-4240 Sequence 4240, Ap
145 27 65.9 98 1 US-08-211-202-111 Sequence 111, App
146 27 65.9 98 2 US-08-665-202-38 Sequence 38, Appl
147 27 65.9 98 4 US-09-315-574-38 Sequence 38, Appl
148 27 65.9 109 3 US-09-240-274-54 Sequence 54, Appl
149 27 65.9 109 3 US-09-240-274-58 Sequence 58, Appl
150 27 65.9 109 3 US-09-240-274-58 Sequence 58, Appl

ALIGNMENTS

US-09-374-454-19
; Sequence 19, Application US/09374454
; Patent No. 6395548
; GENERAL INFORMATION:
; APPLICANT: Lee, Mu-En
; APPLICANT: Maemura, Koji
; APPLICANT: Heieh, Chung-Ming
; TITLE OF INVENTION: METHODS OF MODULATING OF ANGIOGENESIS
; FILE REFERENCE: 05433/037001
; CURRENT APPLICATION NUMBER: US/09/374,454
; CURRENT FILING DATE: 1999-08-13
; EARLIER APPLICATION NUMBER: US 60/096,515
; EARLIER FILING DATE: 1998-08-14
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 602

; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-374-454-19
Query Match 87.8%; Score 36; DB 4; Length 602;
Best Local Similarity 85.7%; Pred. No. 98;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
Db 135 GSNYRPS 141

RESULT 2
US-09-489-039A-11409
; Sequence 11409, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 11409
; LENGTH: 157
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-11409
Query Match 80.5%; Score 33; DB 4; Length 157;
Best Local Similarity 83.3%; Pred. No. 95;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNYRP 6
Db 129 GNNYQP 134

RESULT 3
US-09-489-039A-11252
; Sequence 11252, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 11252
; LENGTH: 324
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-11252
Query Match 80.5%; Score 33; DB 4; Length 324;
Best Local Similarity 71.4%; Pred. No. 1.8e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
Db 153 GNDYRPT 159

RESULT 4
US-08-345-321-4
```

```

; Sequence 4, Application US/08345321
; Patent No. 5914109
; GENERAL INFORMATION:
; APPLICANT: ZOLLA-PAZNER, Susan
; APPLICANT: GORNY, Mikoslav K.
; TITLE OF INVENTION: HETEROHYBRIDOMAS PRODUCING HUMAN
; TITLE OF INVENTION: MONOCLONAL ANTIBODIES TO HIV-1
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Broadway and Neimark
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/345,321
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/872,675
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Browdy, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: ZOLLA-PAZNER1B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 132 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-345-321-4

Query Match 78.0%; Score 32; DB 2; Length 132;
Best Local Similarity 85.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
DB 70 GNNKRP 76

RESULT 5
US-09-621-976-4449
; Sequence 4449, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 4449
; LENGTH: 159
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-621-976-4449

Query Match 78.0%; Score 32; DB 4; Length 159;

```

```

Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
DB 37 GNNRPS 43

RESULT 6
US-09-107-532A-7214
; Sequence 7214, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lytun A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 7214:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 223 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (8) LOCATION 1...223
; SEQUENCE DESCRIPTION: SEQ ID NO: 7214:

US-09-107-532A-7214

Query Match 78.0%; Score 32; DB 4; Length 223;
Best Local Similarity 83.3%; Pred. No. 1.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNYRPS 6
DB 141 GDNYP 146

RESULT 7
US-08-035-392-2
; Sequence 2, Application US/08035392
; Patent No. 5484732
; GENERAL INFORMATION:

```

APPLICANT: Rosenfeld, M. G.
APPLICANT: Andersen, B.
TITLE OF INVENTION: TRANSCRIPTION FACTOR FOR THE REGULATION
OF THE DEVELOPMENT OF SKIN AND HAIR
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Spensley Horn Jubas & Lubitz
STREET: 1880 Century Park East - Suite 500
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90067
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/035,392
FILING DATE: 19930322
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Howells, Stacy L.
REGISTRATION NUMBER: 34,842
REFERENCE/DOCKET NUMBER: PD-2607
TELEPHONE: (619) 455-5100
TELEFAX: (619) 455-5110
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 348 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-035-392-2

Query Match 78.0%; Score 32; DB 1; Length 348;
Best Local Similarity 85.7%; Pred. No. 2.9e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
DB 297 GNNSRPS 303

RESULT 8
US-08-504-511A-2
Sequence 2, Application US/08504511A
Patent No. 5561224
GENERAL INFORMATION:
APPLICANT: Rosenfeld, M. G.
APPLICANT: Andersen, B.
TITLE OF INVENTION: TRANSCRIPTION FACTOR FOR THE REGULATION
OF THE DEVELOPMENT OF SKIN AND HAIR
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: California
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/504,511A
FILING DATE: 20-JUL-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Howells, Stacy L.

REGISTRATION NUMBER: 34,842
REFERENCE/DOCKET NUMBER: 07340/002002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 678-5070
TELEFAX: (619) 678-5099
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 348 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-504-511A-2

Query Match 78.0%; Score 32; DB 1; Length 348;
Best Local Similarity 85.7%; Pred. No. 2.9e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
DB 297 GNNSRPS 303

RESULT 9
US-08-035-392-4
Sequence 4, Application US/08035392
Patent No. 5484732
GENERAL INFORMATION:
APPLICANT: Rosenfeld, M. G.
APPLICANT: Andersen, B.
TITLE OF INVENTION: TRANSCRIPTION FACTOR FOR THE REGULATION
OF THE DEVELOPMENT OF SKIN AND HAIR
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Spensley Horn Jubas & Lubitz
STREET: 1880 Century Park East - Suite 500
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90067
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/035,392
FILING DATE: 19930322
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Howells, Stacy L.
REGISTRATION NUMBER: 34,842
REFERENCE/DOCKET NUMBER: PD-2607
TELEPHONE: (619) 455-5100
TELEFAX: (619) 455-5110
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 430 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-035-392-4

Query Match 78.0%; Score 32; DB 1; Length 430;
Best Local Similarity 85.7%; Pred. No. 3.5e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
DB 379 GNNSRPS 385

RESULT 10

US-08-504-511A-4
; Sequence 4, Application US/08504511A
; Patent No. 5561224
; GENERAL INFORMATION:

; APPLICANT: Rosenfeld, M. G.
; APPLICANT: Andersen, B.
; TITLE OF INVENTION: TRANSCRIPTION FACTOR FOR THE REGULATION
; TITLE OF INVENTION: OF THE DEVELOPMENT OF SKIN AND HAIR
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: California
; COUNTRY: USA
; ZIP: 92037

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/504,511A
; FILING DATE: 20-JUL-1995

; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Howells, Stacy L.
; REGISTRATION NUMBER: 34,842
; REFERENCE/DOCKET NUMBER: 07340/002002

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 678-5070
; TELEFAX: (619) 678-5099
; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 430 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-504-511A-4

Query Match 78.0%; Score 32; DB 1; Length 430;
Best Local Similarity 85.7%; Pred. No. 3.5e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
||| |||
DB 379 GNNSRPS 385

RESULT 11
US-09-540-236-3486
; Sequence 3486, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:

; APPLICANT: Gary L. Bretton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRHALIS
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 3486

; LENGTH: 442
; TYPE: PRT
; ORGANISM: M.catarrhalis
US-09-540-236-3486

Query Match 78.0%; Score 32; DB 4; Length 442;
Best Local Similarity 85.7%; Pred. No. 3.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
||| |||

DB 213 GNNSRPS 219

RESULT 12

US-09-835-654-2
; Sequence 2, Application US/09835654
; Patent No. 6646182
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; TITLE OF INVENTION: Mre11 Orthologue and Uses Thereof
; FILE REFERENCE: 1264
; CURRENT APPLICATION NUMBER: US/09/835,654
; CURRENT FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: US 60/198,570
; PRIOR FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 552
; TYPE: PRT
; ORGANISM: Zea mays
US-09-835-654-2

Query Match 78.0%; Score 32; DB 4; Length 552;
Best Local Similarity 71.4%; Pred. No. 4.4e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
||| |||
DB 296 GNNYRPT 302

RESULT 13

US-09-480-921B-8
; Sequence 8, Application US/09480921B
; Patent No. 6387637
; GENERAL INFORMATION:
; APPLICANT: Levin, Joshua Z.
; APPLICANT: Budziszewski, Gregory J.
; APPLICANT: Potter, Sharon L.
; APPLICANT: Wegrich, Lynette M.
; TITLE OF INVENTION: Herbicide Target Genes and Methods
; FILE REFERENCE: PB/5-30780A
; CURRENT APPLICATION NUMBER: US/09/480,921B
; CURRENT FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 720
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-480-921B-8

Query Match 78.0%; Score 32; DB 4; Length 720;
Best Local Similarity 71.4%; Pred. No. 5.6e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
||| |||
DB 296 GNNYRPT 302

RESULT 14

US-09-976-594-893
; Sequence 893, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12

; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 893
; LENGTH: 1115
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6673549 2743049CD1
US-09-976-594-893

Query Match 78.0%; Score 32; DB 4; Length 1115;
Best Local Similarity 83.3%; Pred. No. 8.4e+02; Indels 0; Gaps 0;
Matches 5; Conservative 1; Mismatches 0;

Qy 1 GNNYRP 6
Db 955 GNNYRP 960
|:|||||

RESULT 15
US-08-672-345C-38
; Sequence 38, Application US/08672345C
; Patent No. 5948658
; GENERAL INFORMATION:
; APPLICANT: Landry Donald, W.
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY
; NUMBER OF SEQUENCES: 108
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper and Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/672,345C
; FILING DATE: 24-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/51400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-672-345C-38

Query Match 75.6%; Score 31; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 3e+05; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0;

Qy 2 NNYRP 6
Db 2 NNYRP 6
|||||

RESULT 16
US-08-672-345C-74

; Sequence 74, Application US/08672345C
; Patent No. 5948658
; GENERAL INFORMATION:
; APPLICANT: Landry Donald, W.
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY
; NUMBER OF SEQUENCES: 108
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper and Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/672,345C
; FILING DATE: 24-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/51400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 74:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-672-345C-74

Query Match 75.6%; Score 31; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 3e+05; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0;

Qy 2 NNYRP 6
Db 2 NNYRP 6
|||||

RESULT 17
US-09-214-095D-38
; Sequence 38, Application US/09214095D
; Patent No. 6280987
; GENERAL INFORMATION:
; APPLICANT: Landry, Donald
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY
; FILE REFERENCE: 51400-A-PCT-US
; CURRENT APPLICATION NUMBER: US/09/214,095D
; CURRENT FILING DATE: 1993-07-19
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 38
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
US-09-214-095D-38

Query Match 75.6%; Score 31; DB 3; Length 7;
Best Local Similarity 100.0%; Pred. No. 3e+05; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0;

Qy 2 NNYRP 6
Db 2 NNYRP 6
|||||

RESULT 18
US-09-214-095D-74
; Sequence 74, Application US/09214095D
; Patent No. 6280987
; GENERAL INFORMATION:
; APPLICANT: Landry, Donald
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY
; FILE REFERENCE: 51400-A-PCT-US
; CURRENT APPLICATION NUMBER: US/09/214,095D
; CURRENT FILING DATE: 1999-07-19
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 74
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
; FEATURE:
; NAME/KEY: CHAIN
; LOCATION: (1)..(1)
; OTHER INFORMATION: X at position 1 represent any amino acid
US-09-214-095D-74

Query Match 75.6%; Score 31; DB 3; Length 7;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 NNYP 6
|||
Db 2 NNYP 6

RESULT 19
US-09-621-976-4090
; Sequence 4090, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 4090
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -15...-1
US-09-621-976-4090

Query Match 75.6%; Score 31; DB 4; Length 74;
Best Local Similarity 71.4%; Pred. No. 1.1e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
|||
Db 50 GNEYQPS 56

RESULT 20
US-09-621-976-5643
; Sequence 5643, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.

; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 5643
; LENGTH: 75
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -15...-1
US-09-621-976-5643

Query Match 75.6%; Score 31; DB 4; Length 75;
Best Local Similarity 71.4%; Pred. No. 1.1e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
|||
Db 50 GNEYQPS 56

RESULT 21
US-08-672-345C-3
; Sequence 3, Application US/08672345C
; Patent No. 5948658
; GENERAL INFORMATION:
; APPLICANT: Landry Donald, W.
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY
; NUMBER OF SEQUENCES: 108
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper and Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/672,345C
; FILING DATE: 24-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/51400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 109 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-672-345C-3

Query Match 75.6%; Score 31; DB 2; Length 109;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 NNYP 6
|||
Db 52 NNYP 56

RESULT 22

US-08-672-345C-93
; Sequence 93, Application US/08672345C
; Patent No. 5948658
; GENERAL INFORMATION:
; APPLICANT: Landry Donald, W.
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY
; NUMBER OF SEQUENCES: 108
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper and Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/672,345C
; FILING DATE: 24-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/51400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 93:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 109 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-672-345C-93

Query Match 75.6%; Score 31; DB 2; Length 109;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 NNYRP 6
Db 52 NNYRP 56

RESULT 23
US-09-214-095D-3
; Sequence 3, Application US/09214095D
; Patent No. 6280987
; GENERAL INFORMATION:
; APPLICANT: Landry, Donald
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY
; FILE REFERENCE: 51400-A-PCT-US
; CURRENT APPLICATION NUMBER: US/09/214,095D
; CURRENT FILING DATE: 1999-07-19
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 109
; TYPE: PRT
; ORGANISM: Murinae gen. Sp.
US-09-214-095D-3

Query Match 75.6%; Score 31; DB 3; Length 109;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 NNYRP 6
Db 52 NNYRP 56

RESULT 24
US-09-214-095D-121
; Sequence 121, Application US/09214095D
; Patent No. 6280987
; GENERAL INFORMATION:
; APPLICANT: Landry, Donald
; TITLE OF INVENTION: ANTI-COCAINE CATALYTIC ANTIBODY
; FILE REFERENCE: 51400-A-PCT-US
; CURRENT APPLICATION NUMBER: US/09/214,095D
; CURRENT FILING DATE: 1999-07-19
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 121
; LENGTH: 109
; TYPE: PRT
; ORGANISM: Murine
US-09-214-095D-121

Query Match 75.6%; Score 31; DB 3; Length 109;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 NNYRP 6
Db 52 NNYRP 56

RESULT 25
US-09-489-039A-10928
; Sequence 10928, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10928
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10928

Query Match 75.6%; Score 31; DB 4; Length 330;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 NNYRP 6
Db 188 NNYRP 192

Search completed: September 24, 2004, 02:19:58
Job time : 14.8868 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 01:54:43 ; Search time 40.5472 Seconds
(without alignments)
55.513 Million cell updates/sec

Title: US-10-088-639A-2_COPY_49_55

Perfect score: 41

Sequence: 1 GNNYRPS 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1349238 seqs, 321558718 residues

Total number of hits satisfying chosen parameters: 1349238

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Lasting first 150 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US03_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/CTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41	100.0	7	16	US-10-327-598-656
2	36	87.8	201	9	US-09-764-864-1561
3	36	87.8	588	12	US-10-467-721-8
4	36	87.8	591	9	US-09-764-864-1139
5	36	87.8	599	12	US-10-467-721-6
6	36	87.8	602	12	US-10-260-708-78
7	36	87.8	602	14	US-10-121-235-19
8	36	87.8	622	12	US-10-467-721-4
9	36	87.8	636	12	US-10-467-721-2
10	35	85.4	251	10	US-09-880-748-1555
11	35	85.4	251	12	US-10-293-418-1555
12	34	82.9	919	15	US-10-369-493-1435
13	34	82.9	919	16	US-10-149-310-238
14	33	80.5	110	16	US-10-723-434-53
15	33	80.5	243	14	US-10-322-673-51
					Sequence 656, App
					Sequence 1561, App
					Sequence 8, Appli
					Sequence 1139, App
					Sequence 6, Appli
					Sequence 78, Appl
					Sequence 19, Appl
					Sequence 4, Appli
					Sequence 2, Appli
					Sequence 1555, App
					Sequence 1555, App
					Sequence 1435, App
					Sequence 238, App
					Sequence 53, Appli
					Sequence 51, Appli

16	33	80.5	247	10	US-09-880-748-1993	Sequence 1993, Ap
17	33	80.5	247	10	US-09-880-748-1998	Sequence 1998, Ap
18	33	80.5	247	12	US-10-293-418-1993	Sequence 1993, Ap
19	33	80.5	247	12	US-10-293-418-1998	Sequence 1998, Ap
20	33	80.5	247	16	US-10-779-461-44	Sequence 44, Appl
21	33	80.5	248	10	US-09-880-748-1976	Sequence 1976, Ap
22	33	80.5	248	10	US-09-880-748-2000	Sequence 2000, Ap
23	33	80.5	248	12	US-10-293-418-1976	Sequence 1976, Ap
24	33	80.5	248	12	US-10-293-418-2000	Sequence 2000, Ap
25	33	80.5	254	10	US-09-880-748-1958	Sequence 1958, Ap
26	33	80.5	254	12	US-10-293-418-1958	Sequence 1958, Ap
27	33	80.5	258	10	US-09-880-748-2083	Sequence 2083, Ap
28	33	80.5	258	12	US-10-293-418-2083	Sequence 2083, Ap
29	33	80.5	832	15	US-10-094-749-1919	Sequence 1919, Ap
30	33	80.5	861	16	US-10-451-168-75	Sequence 75, Appl
31	33	80.5	922	16	US-10-451-168-77	Sequence 77, Appl
32	33	80.5	1274	9	US-09-746-491-12	Sequence 12, Appl
33	32	78.0	7	14	US-10-203-754A-14	Sequence 14, Appl
34	32	78.0	7	16	US-10-327-598-626	Sequence 626, App
35	32	78.0	7	16	US-10-327-598-698	Sequence 698, App
36	32	78.0	58	12	US-10-424-599-249492	Sequence 249492,
37	32	78.0	116	14	US-10-203-754A-20	Sequence 20, Appl
38	32	78.0	124	12	US-10-424-599-195801	Sequence 195801,
39	32	78.0	139	16	US-10-327-598-1078	Sequence 1078, Ap
40	32	78.0	148	16	US-10-767-701-49911	Sequence 49911, A
41	32	78.0	177	12	US-10-424-599-240460	Sequence 240460,
42	32	78.0	182	16	US-10-767-701-55203	Sequence 55203, A
43	32	78.0	190	12	US-10-424-599-277708	Sequence 277708,
44	32	78.0	199	16	US-10-437-963-164100	Sequence 164100,
45	32	78.0	221	12	US-10-424-599-224918	Sequence 224918,
46	32	78.0	238	16	US-10-779-461-1	Sequence 1, Appli
47	32	78.0	246	10	US-09-880-748-1847	Sequence 1847, Ap
48	32	78.0	246	12	US-10-293-418-1847	Sequence 1847, Ap
49	32	78.0	247	10	US-09-880-748-1978	Sequence 1978, Ap
50	32	78.0	247	12	US-10-293-418-1978	Sequence 1978, Ap
51	32	78.0	248	10	US-09-880-748-1031	Sequence 1031, Ap
52	32	78.0	248	10	US-09-880-748-1237	Sequence 1237, Ap
53	32	78.0	248	12	US-10-293-418-1031	Sequence 1031, Ap
54	32	78.0	248	12	US-10-293-418-1237	Sequence 1237, Ap
55	32	78.0	249	10	US-09-880-748-787	Sequence 787, App
56	32	78.0	249	10	US-09-880-748-1142	Sequence 1142, Ap
57	32	78.0	249	10	US-09-880-748-1984	Sequence 1984, Ap
58	32	78.0	249	10	US-09-880-748-1997	Sequence 1997, App
59	32	78.0	249	12	US-10-293-418-787	Sequence 787, App
60	32	78.0	249	12	US-10-293-418-1142	Sequence 1142, Ap
61	32	78.0	249	12	US-10-293-418-1984	Sequence 1984, Ap
62	32	78.0	249	12	US-10-293-418-1997	Sequence 1997, Ap
63	32	78.0	250	16	US-10-779-461-28	Sequence 28, Appli
64	32	78.0	251	10	US-09-880-748-6	Sequence 6, Appli
65	32	78.0	251	10	US-09-880-748-1225	Sequence 1225, Ap
66	32	78.0	251	10	US-09-880-748-1231	Sequence 1231, Ap
67	32	78.0	251	10	US-09-880-748-1712	Sequence 1712, Ap
68	32	78.0	251	12	US-10-293-418-6	Sequence 6, Appli
69	32	78.0	251	12	US-10-293-418-1225	Sequence 1225, Ap
70	32	78.0	251	12	US-10-293-418-1231	Sequence 1231, Ap
71	32	78.0	251	12	US-10-293-418-1712	Sequence 1712, Ap
72	32	78.0	251	14	US-10-120-414-75	Sequence 75, Appli
73	32	78.0	252	10	US-09-880-748-1144	Sequence 1144, Ap
74	32	78.0	252	10	US-09-880-748-1201	Sequence 1201, Ap
75	32	78.0	252	10	US-09-880-748-1519	Sequence 1519, Ap
76	32	78.0	252	10	US-09-880-748-1575	Sequence 1575, Ap
77	32	78.0	252	12	US-10-293-418-1144	Sequence 1144, Ap
78	32	78.0	252	12	US-10-293-418-1201	Sequence 1201, Ap
79	32	78.0	252	12	US-10-293-418-1519	Sequence 1519, Ap
80	32	78.0	252	12	US-10-293-418-1575	Sequence 1575, Ap
81	32	78.0	253	10	US-09-880-748-1482	Sequence 1482, Ap
82	32	78.0	253	12	US-10-293-418-1482	Sequence 1482, Ap
83	32	78.0	255	12	US-09-880-748-1862	Sequence 1862, Ap
84	32	78.0	255	12	US-10-293-418-1862	Sequence 1862, Ap
85	32	78.0	260	10	US-09-880-748-1282	Sequence 1282, Ap
86	32	78.0	260	12	US-10-293-418-1282	Sequence 1282, Ap
87	32	78.0	288	12	US-10-406-830-3	Sequence 3, Appli
88	32	78.0	306	12	US-10-424-599-179002	Sequence 179002,

89 32 78.0 352 16 US-10-437-963-120350
90 32 78.0 357 11 US-09-764-875-792
91 32 78.0 395 12 US-10-424-599-200998
92 32 78.0 417 12 US-10-362-881-39
93 32 78.0 440 9 US-09-764-853-594
94 32 78.0 440 11 US-09-764-875-1102
95 32 78.0 442 11 US-09-764-875-809
96 32 78.0 449 9 US-09-764-853-577
97 32 78.0 456 16 US-10-437-963-167355
98 32 78.0 490 12 US-10-425-114-53913
99 32 78.0 497 12 US-10-425-114-64871
100 32 78.0 533 9 US-10-424-599-269344
101 32 78.0 552 9 US-09-835-654-2
102 32 78.0 552 12 US-10-650-108-2
103 32 78.0 552 16 US-10-437-963-109-2
104 32 78.0 554 16 US-10-437-963-184295
105 32 78.0 579 16 US-10-437-963-167357
106 32 78.0 628 14 US-10-213-990-36
107 32 78.0 649 15 US-10-259-194A-250
108 32 78.0 720 13 US-10-047-412A-8
109 31 75.6 7 10 US-09-940-727B-38
110 31 75.6 7 16 US-09-940-727B-74
111 31 75.6 7 16 US-10-688-925-41
112 31 75.6 7 16 US-10-327-598-625
113 31 75.6 7 16 US-10-327-598-649
114 31 75.6 7 16 US-10-327-598-686
115 31 75.6 85 12 US-10-424-599-206386
116 31 75.6 92 9 US-09-731-872-370
117 31 75.6 92 10 US-09-876-997-370
118 31 75.6 92 16 US-10-258-662-18
119 31 75.6 94 16 US-10-767-701-37127
120 31 75.6 98 12 US-10-424-599-187869
121 31 75.6 109 10 US-09-940-727B-3
122 31 75.6 109 16 US-09-940-727B-121
123 31 75.6 111 16 US-10-688-925-12
124 31 75.6 111 16 US-10-688-925-24
125 31 75.6 126 10 US-09-764-891-4786
126 31 75.6 126 14 US-10-091-572-341
127 31 75.6 128 16 US-10-437-963-120634
128 31 75.6 148 12 US-10-282-122A-49010
129 31 75.6 149 9 US-09-738-626-6793
130 31 75.6 151 16 US-10-466-531-85
131 31 75.6 169 16 US-10-437-963-197676
132 31 75.6 170 16 US-10-767-701-53508
133 31 75.6 250 10 US-09-880-748-2093
134 31 75.6 250 12 US-09-880-748-2097
135 31 75.6 250 12 US-10-293-418-2093
136 31 75.6 250 12 US-10-293-418-2097
137 31 75.6 253 10 US-09-880-748-1895
138 31 75.6 253 10 US-09-880-748-2098
139 31 75.6 253 12 US-09-880-748-2101
140 31 75.6 253 12 US-10-293-418-1895
141 31 75.6 253 12 US-10-293-418-2098
142 31 75.6 253 12 US-10-293-418-2101
143 31 75.6 256 10 US-09-880-748-1967
144 31 75.6 256 12 US-10-293-418-1967
145 31 75.6 258 16 US-10-389-566-1045
146 31 75.6 258 16 US-10-688-925-8
147 31 75.6 258 16 US-10-688-925-20
148 31 75.6 282 12 US-10-282-122A-57410
149 31 75.6 303 15 US-10-369-493-13986
150 31 75.6 319 12 US-10-282-122A-47606

ALIGNMENTS

RESULT 1

US-10-327-598-656
; Sequence 656, Application US/10327598
; Publication No. US20040181039A1
; GENERAL INFORMATION:
; APPLICANT: Krahn, Eugene

Sequence 120350,
Sequence 792, App
Sequence 200998,
Sequence 39, Appl
Sequence 594, App
Sequence 1102, App
Sequence 809, App
Sequence 577, App
Sequence 167355,
Sequence 53913, A
Sequence 64871, A
Sequence 269344,
Sequence 2, Appl
Sequence 2, Appl
Sequence 2, Appl
Sequence 184295,
Sequence 167357,
Sequence 36, Appl
Sequence 250, App
Sequence 8, Appl
Sequence 38, Appl
Sequence 41, Appl
Sequence 625, App
Sequence 649, App
Sequence 686, App
Sequence 206386,
Sequence 370, App
Sequence 370, App
Sequence 18, Appl
Sequence 37127, A
Sequence 187869,
Sequence 3, Appl
Sequence 12, App
Sequence 24, Appl
Sequence 4786, App
Sequence 341, App
Sequence 120634,
Sequence 49010, A
Sequence 6793, App
Sequence 85, Appl
Sequence 197676,
Sequence 53508, A
Sequence 2093, App
Sequence 2097, App
Sequence 2097, App
Sequence 1895, App
Sequence 2098, App
Sequence 2101, App
Sequence 1895, App
Sequence 2098, App
Sequence 2101, App
Sequence 1967, App
Sequence 1967, App
Sequence 1045, App
Sequence 8, Appl
Sequence 20, Appl
Sequence 57410, A
Sequence 13986, A
Sequence 47606, A

; APPLICANT: Guo, Honliang
; APPLICANT: Aiyappa, Ashok
; APPLICANT: Lawton, Robert
; TITLE OF INVENTION: Canine Immunoglobulin Variable Domains, Caninized Antibodies, and
; TITLE OF INVENTION: for Making and Using Them
; FILE REFERENCE: 01-799-A
; CURRENT APPLICATION NUMBER: US/10/327,598
; CURRENT FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: US 60/344,874
; PRIOR FILING DATE: 2001-12-21
; NUMBER OF SEQ ID NOS: 1139
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 656
; LENGTH: 7
; TYPE: PRT
; ORGANISM: canis familiaris;
US-10-327-598-656

Query Match 100.0%; Score 41; DB 16; Length 7;

Best Local Similarity 100.0%; Pred. No. 1.2e+06; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;

QY 1 GNNYRPS 7
| | | | |
Db 1 GNNYRPS 7

RESULT 2

US-09-764-864-1561
; Sequence 1561, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 1561
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (176)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (195)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1561

Query Match 87.8%; Score 36; DB 9; Length 201;

Best Local Similarity 85.7%; Pred. No. 53;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
| : | | | | |
Db 124 GSNYRPS 130

RESULT 3

US-10-467-721-8
; Sequence 8, Application US/10467721
; Publication No. US20040058366A1
; GENERAL INFORMATION:
; APPLICANT: JAPAN SCIENCE AND TECHNOLOGY CORPORATION
; TITLE OF INVENTION: Bmal2, novel clock genes
; FILE REFERENCE: A011-15PCT
; CURRENT APPLICATION NUMBER: US/10/467,721
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: JP 2001/35743

; PRIOR FILING DATE: 2001-02-13
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 588
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-467-721-8

Query Match 87.8%; Score 36; DB 12; Length 588;
Best Local Similarity 85.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
Db 121 GSNYRPS 127
|:|||||

RESULT 4

US-09-764-864-1139
; Sequence 1139, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17

; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1139
; LENGTH: 591
; TYPE: PRT
; ORGANISM: Homo sapiens

; FEATURE:
; NAME/KEY: SITE
; LOCATION: (338)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; NAME/KEY: SITE
; LOCATION: (376)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; NAME/KEY: SITE
; LOCATION: (465)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; NAME/KEY: SITE
; LOCATION: (495)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; NAME/KEY: SITE
; LOCATION: (491)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; NAME/KEY: SITE
; LOCATION: (491)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; NAME/KEY: SITE
; LOCATION: (491)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; NAME/KEY: SITE
; LOCATION: (491)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; NAME/KEY: SITE
; LOCATION: (491)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; NAME/KEY: SITE
; LOCATION: (491)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; NAME/KEY: SITE
; LOCATION: (491)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; NAME/KEY: SITE
; LOCATION: (491)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; NAME/KEY: SITE
; LOCATION: (491)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; NAME/KEY: SITE
; LOCATION: (491)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; NAME/KEY: SITE
; LOCATION: (491)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; NAME/KEY: SITE
; LOCATION: (491)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; NAME/KEY: SITE
; LOCATION: (491)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; NAME/KEY: SITE
; LOCATION: (491)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; NAME/KEY: SITE
; LOCATION: (491)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; NAME/KEY: SITE
; LOCATION: (491)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; NAME/KEY: SITE
; LOCATION: (491)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; NAME/KEY: SITE
; LOCATION: (491)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; NAME/KEY: SITE
; LOCATION: (491)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; PRIOR FILING DATE: 2001-02-13
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 599
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-467-721-6

Query Match 87.8%; Score 36; DB 12; Length 599;
Best Local Similarity 85.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
Db 132 GSNYRPS 138
|:|||||

RESULT 6

US-10-260-708-78
; Sequence 78, Application US/10260708
; Publication No. US20040063101A1
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew
; APPLICANT: Lee, Sang-Yull
; APPLICANT: Old, Lloyd

; TITLE OF INVENTION: Human Sarcoma-Associated Antigens
; FILE REFERENCE: L00461/70138
; CURRENT APPLICATION NUMBER: US/10/260,708
; CURRENT FILING DATE: 2002-09-30
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 78
; LENGTH: 602
; TYPE: PRT
; ORGANISM: homo sapiens

US-10-260-708-78

US-10-260-708-78

US-10-260-708-78

US-10-260-708-78

US-10-260-708-78

US-10-260-708-78

US-10-260-708-78

US-10-260-708-78

US-10-260-708-78

US-10-260-708-78

US-10-260-708-78

US-10-260-708-78

US-10-260-708-78

US-10-260-708-78

US-10-260-708-78

US-10-260-708-78

US-10-260-708-78

US-10-260-708-78

US-10-260-708-78

US-10-260-708-78

US-10-260-708-78

US-10-260-708-78

US-10-260-708-78

US-10-260-708-78

US-10-260-708-78

US-10-260-708-78

US-10-260-708-78

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNYRPS 7

Db 135 GSNYRPS 141

RESULT 8

US-10-467-721-4
; Sequence 4, Application US/10467721
; Publication No. US20040058366A1
; GENERAL INFORMATION:
; APPLICANT: JAPAN SCIENCE AND TECHNOLOGY CORPORATION
; TITLE OF INVENTION: Bmal2, novel clock genes
; FILE REFERENCE: A011-15PCT
; CURRENT APPLICATION NUMBER: US/10/467,721
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: JP 2001/35743
; PRIOR FILING DATE: 2001-02-13
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 622
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-467-721-4

Query Match 87.8%; Score 36; DB 12; Length 622;
Best Local Similarity 85.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNYRPS 7

Db 155 GSNYRPS 161

RESULT 9

US-10-467-721-2
; Sequence 2, Application US/10467721
; Publication No. US20040058366A1
; GENERAL INFORMATION:
; APPLICANT: JAPAN SCIENCE AND TECHNOLOGY CORPORATION
; TITLE OF INVENTION: Bmal2, novel clock genes
; FILE REFERENCE: A011-15PCT
; CURRENT APPLICATION NUMBER: US/10/467,721
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: JP 2001/35743
; PRIOR FILING DATE: 2001-02-13
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 636
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-467-721-2

Query Match 87.8%; Score 36; DB 12; Length 636;
Best Local Similarity 85.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNYRPS 7

Db 169 GSNYRPS 175

RESULT 10

US-09-880-748-1555
; Sequence 1555, Application US/09880748
; Publication No. US20030059937A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind BlyS
; FILE REFERENCE: PF523

; CURRENT APPLICATION NUMBER: US/09/880,748

; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 3239
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1555
; LENGTH: 251
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-748-1555

Query Match 85.4%; Score 35; DB 10; Length 251;
Best Local Similarity 85.7%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNNYRPS 7

Db 191 GKNYRPS 197

RESULT 11

US-10-293-418-1555
; Sequence 1555, Application US/10293418
; Publication No. US2003022396A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind BlyS
; FILE REFERENCE: PF523P2
; CURRENT APPLICATION NUMBER: US/10/293,418
; CURRENT FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: 60/331,469
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/340,817
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 09/880,748
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1555
; LENGTH: 251
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-418-1555

Query Match 85.4%; Score 35; DB 12; Length 251;
Best Local Similarity 85.7%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNNYRPS 7

Db 191 GKNYRPS 197

RESULT 12

US-10-369-493-1435

; Sequence 1435, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 1435
; LENGTH: 919
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-369-493-1435

Query Match 82.9%; Score 34; DB 15; Length 919;
Best Local Similarity 85.7%; Pred. No. 5.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
Db 130 GNNYLPs 136

RESULT 13
US-10-149-310-238
; Sequence 238, Application US/10149310
; Publication No. US20040077039A1
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas
; APPLICANT: Madden, Kevin T.
; APPLICANT: Maxon, Mary
; APPLICANT: Sherman, Amir
; TITLE OF INVENTION: Modulation of Secondary Metabolite Production by
; FILE REFERENCE: 14184-019US1
; CURRENT APPLICATION NUMBER: US/10/149,310
; CURRENT FILING DATE: 2003-02-19
; PRIOR APPLICATION NUMBER: PCT/US01/29288
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: US 60/233,564
; PRIOR FILING DATE: 2000-09-19
; NUMBER OF SEQ ID NOS: 308
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 238
; LENGTH: 919
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-149-310-238

Query Match 82.9%; Score 34; DB 15; Length 919;
Best Local Similarity 85.7%; Pred. No. 5.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
Db 130 GNNYLPs 136

RESULT 14
US-10-723-434-53
; Sequence 53, Application US/10723434
; Publication No. US20040133357A1
; GENERAL INFORMATION:
; APPLICANT: Zhong, Pingyu
; APPLICANT: Luo, Peizhi

; APPLICANT: Wang, Kevin C.
; APPLICANT: Li, Yan
; TITLE OF INVENTION: HUMANIZED ANTIBODIES AGAINST VASCULAR ENDOTHELIAL GROWTH FACTOR
; FILE REFERENCE: 26050-709.501
; CURRENT APPLICATION NUMBER: US/10/723,434
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: US 60/284,407
; PRIOR FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: US 10/125,687
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: US 10/153,176
; PRIOR FILING DATE: 2002-05-20
; PRIOR APPLICATION NUMBER: US 10/443,134
; PRIOR FILING DATE: 2003-05-20
; NUMBER OF SEQ ID NOS: 156
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 53
; LENGTH: 110
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: VL
US-10-723-434-53

Query Match 80.5%; Score 33; DB 16; Length 110;
Best Local Similarity 85.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
Db 51 GNNQRPs 57

RESULT 15
US-10-322-673-51
; Sequence 51, Application US/10322673
; Publication No. US20030180296A1
; GENERAL INFORMATION:
; APPLICANT: Salcedo et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind to TRAIL
; FILE REFERENCE: PFS95
; CURRENT APPLICATION NUMBER: US/10/322,673
; CURRENT FILING DATE: 2002-12-19
; PRIOR APPLICATION NUMBER: 60/341,237
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/369,877
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/384,828
; PRIOR FILING DATE: 2002-06-04
; PRIOR APPLICATION NUMBER: 60/396,591
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/403,370
; PRIOR FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: 60/425,737
; PRIOR FILING DATE: 2002-11-13
; NUMBER OF SEQ ID NOS: 72
; SEQ ID NO 51
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: CM087C06 scFv
US-10-322-673-51

Query Match 80.5%; Score 33; DB 14; Length 243;
Best Local Similarity 85.7%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
Db 183 GNNQRPs 189

RESULT 16
US-09-880-748-1993
; Sequence 1993, Application US/09880748
; Publication No. US2003005937A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523
; CURRENT APPLICATION NUMBER: US/09/880,748
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 3239
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1993
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-748-1993

Query Match 80.5%; Score 33; DB 10; Length 247;
Best Local Similarity 85.7%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
||| |||
Db 187 GNNQPS 193

RESULT 17
US-09-880-748-1998
; Sequence 1998, Application US/09880748
; Publication No. US2003005937A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523
; CURRENT APPLICATION NUMBER: US/09/880,748
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 3239
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1998
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-748-1998

Query Match 80.5%; Score 33; DB 10; Length 247;
Best Local Similarity 85.7%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
||| |||

Db 187 GNNQPS 193
RESULT 18
US-10-293-418-1993
; Sequence 1993, Application US/10293418
; Publication No. US20030223996A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523P2
; CURRENT APPLICATION NUMBER: US/10/293,418
; CURRENT FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: 60/331,469
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/340,817
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 09/880,748
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/212,210
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1993
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-418-1993

Query Match 80.5%; Score 33; DB 12; Length 247;
Best Local Similarity 85.7%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
||| |||
Db 187 GNNQPS 193

RESULT 19
US-10-293-418-1998
; Sequence 1998, Application US/10293418
; Publication No. US20030223996A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523P2
; CURRENT APPLICATION NUMBER: US/10/293,418
; CURRENT FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: 60/331,469
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/340,817
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 09/880,748
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/212,210
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1998

; LENGTH: 247
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-418-1998

Query Match 80.5%; Score 33; DB 12; Length 247;
Best Local Similarity 85.7%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
Db 187 GNNQRPS 193

RESULT 20

US-10-779-461-44
; Sequence 44, Application US/10779461
; Publication No. US20040166344A1
; GENERAL INFORMATION:
; APPLICANT: Morton, Philip A
; TITLE OF INVENTION: ANTIBODIES TO C-MET FOR THE TREATMENT OF CANCERS
; FILE REFERENCE: 00980/1
; CURRENT APPLICATION NUMBER: US/10/779,461
; CURRENT FILING DATE: 2004-02-13
; PRIOR APPLICATION NUMBER: 60/447,073
; PRIOR FILING DATE: 2003-02-13
; NUMBER OF SEQ ID NOS: 161
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 44
; LENGTH: 247
; TYPE: PRT
; ORGANISM: artificial
; FEATURE:
; OTHER INFORMATION: phage display generated human antibody
US-10-779-461-44

Query Match 80.5%; Score 33; DB 16; Length 247;
Best Local Similarity 85.7%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
Db 188 GNNQRPS 194

RESULT 21

US-09-880-748-1976
; Sequence 1976, Application US/09880748
; Publication No. US20030059937A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523
; CURRENT APPLICATION NUMBER: US/09/880,748
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 3239
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1976
; LENGTH: 248
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-748-1976

Query Match 80.5%; Score 33; DB 10; Length 248;
Best Local Similarity 85.7%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
Db 188 GNNQRPS 194

RESULT 22

US-09-880-748-2000
; Sequence 2000, Application US/09880748
; Publication No. US20030059937A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523
; CURRENT APPLICATION NUMBER: US/09/880,748
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 3239
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2000
; LENGTH: 248
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-748-2000

Query Match 80.5%; Score 33; DB 10; Length 248;
Best Local Similarity 85.7%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
Db 188 GNNQRPS 194

RESULT 23

US-10-293-418-1976
; Sequence 1976, Application US/10293418
; Publication No. US20030223996A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523P2
; CURRENT APPLICATION NUMBER: US/10/293,418
; CURRENT FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: 60/331,469
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/340,817
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 09/880,748
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 3247

```
; SEQ ID NO 1976
; LENGTH: 248
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-418-1976

Query Match      80.5%; Score 33; DB 12; Length 248;
Best Local Similarity 85.7%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GNNYRPS 7
      ||| |||
Db      188 GNNQRPS 194

RESULT 24
US-10-293-418-2000
; Sequence 2000, Application US/10293418
; Publication No. US2003022396A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523P2
; CURRENT APPLICATION NUMBER: US/10/293,418
; CURRENT FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: 60/331,469
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/340,817
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 09/880,748
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 2000
; LENGTH: 248
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-418-2000

Query Match      80.5%; Score 33; DB 12; Length 248;
Best Local Similarity 85.7%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GNNYRPS 7
      ||| |||
Db      188 GNNQRPS 194

RESULT 25
US-09-880-748-1958
; Sequence 1958, Application US/09880748
; Publication No. US2003005937A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523
; CURRENT APPLICATION NUMBER: US/09/880,748
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16

; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293,499
; NUMBER OF SEQ ID NOS: 3239
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1958
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-748-1958

Query Match      80.5%; Score 33; DB 10; Length 254;
Best Local Similarity 85.7%; Pred. No. 2.5e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GNNYRPS 7
      ||| |||
Db      193 GNNQRPS 199

Search completed: September 24, 2004, 02:25:09
Job time : 42.5472 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 01:33:11 ; Search time 10.434 Seconds
(without alignments)
64.534 Million cell updates/sec

Title: US-10-088-639A-2_COPY_49_55
Perfect score: 41
Sequence: 1 GNNYRPS 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database :

PIR 78:*
1: piri:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37	90.2	947	2 I49635	mouse Dhml protein
2	35	85.4	222	2 H64422	type II restrictio
3	35	85.4	633	2 S61435	metalloproteinase
4	35	85.4	649	2 S42894	metalloproteinase
5	35	85.4	1182	2 G71607	probable integral
6	34	82.9	919	2 S45889	probable regulator
7	33	80.5	311	2 S23533	hypothetical prote
8	33	80.5	323	2 T00244	sopB protein - Esc
9	33	80.5	342	2 T13113	probable sopB prot
10	33	80.5	491	2 F82843	aminopeptidase A/I
11	33	80.5	520	2 I51556	recombination acti
12	32	78.0	172	2 F86875	peptide methionine
13	32	78.0	340	2 A35009	C 3.4.23.42 thermo
14	32	78.0	348	1 B46216	transcription fact
15	32	78.0	349	2 T49791	hypothetical prote
16	32	78.0	403	1 S35541	transcription fact
17	32	78.0	430	1 A46216	transcription fact
18	32	78.0	431	1 JC2002	transcription fact
19	32	78.0	504	2 H84862	hypothetical prote
20	32	78.0	530	2 F96491	hypothetical prote
21	32	78.0	572	2 F90564	conserved hypotet
22	32	78.0	686	2 AB2362	ribonuclease II [i
23	32	78.0	692	2 AD1857	hypothetical prote
24	32	78.0	720	2 T52564	Mrell protein homo
25	32	78.0	750	2 E84475	probable Athila re
26	32	78.0	935	2 S66306	hypothetical prote
27	31	75.6	105	2 S52680	ribosomal protein
28	31	75.6	108	2 T50348	ribosomal protein
29	31	75.6	184	2 S75635	cobN protein - Syn

30	31	75.6	200	2 A60519	carbonate dehydrat
31	31	75.6	212	2 T27260	hypothetical prote
32	31	75.6	234	2 AD0486	probable exported
33	31	75.6	236	2 AF0209	probable exported
34	31	75.6	258	2 T51689	probable transcrip
35	31	75.6	283	2 T41424	probable diphthine
36	31	75.6	292	2 S89472	protein ZC53.6 [im
37	31	75.6	298	2 S12579	carbonate dehydrat
38	31	75.6	305	1 CRH05	2-dehydro-3-deoxyg
39	31	75.6	309	2 AI0985	2-dehydro-3-deoxyg
40	31	75.6	314	2 AG0485	hypothetical prote
41	31	75.6	318	2 T10238	hypothetical prote
42	31	75.6	327	2 CS5215	2-dehydro-3-deoxyg
43	31	75.6	351	2 A55856	lim protein - Stap
44	31	75.6	365	2 D86470	F21H2.9 protein -
45	31	75.6	377	2 D97102	uncharacterized pr
46	31	75.6	382	2 F91179	ketodeoxygluconoki
47	31	75.6	382	2 G86025	ketodeoxygluconoki
48	31	75.6	382	2 S47747	2-dehydro-3-deoxyg
49	31	75.6	394	2 S78139	translation elonga
50	31	75.6	395	2 I40828	translation elonga
51	31	75.6	395	2 AD1406	translation elonga
52	31	75.6	395	2 AD1782	translation elonga
53	31	75.6	396	2 D60663	translation elonga
54	31	75.6	396	2 A44795	translation elonga
55	31	75.6	396	2 S31151	translation elonga
56	31	75.6	396	2 G87143	elongation factor
57	31	75.6	397	2 S23908	translation elonga
58	31	75.6	397	2 S50138	translation elonga
59	31	75.6	397	2 PC4060	translation elonga
60	31	75.6	397	2 JC5385	translation elonga
61	31	75.6	398	2 F60663	translation elonga
62	31	75.6	398	2 D95173	translation elonga
63	31	75.6	398	2 F98039	translation elonga
64	31	75.6	399	2 A81392	translation elonga
65	31	75.6	427	2 S77207	hypothetical prote
66	31	75.6	429	2 G87636	conserved hypotet
67	31	75.6	437	1 EFBYT	translation elonga
68	31	75.6	511	2 S35417	glucans biosynthes
69	31	75.6	511	2 B90807	periplasmic glucan
70	31	75.6	511	2 F85666	periplasmic glucan
71	31	75.6	513	2 T46143	steroid 22-alpha-h
72	31	75.6	516	2 T02189	hypothetical prote
73	31	75.6	599	2 A57017	RNase L inhibitor
74	31	75.6	599	2 JC6555	ribonuclease L inh
75	31	75.6	599	2 S63672	RNase L inhibitor
76	31	75.6	606	2 T44933	phosphogluconate d
77	31	75.6	752	2 G95094	hypothetical prote
78	31	75.6	752	2 E97962	Arp dependent prot
79	31	75.6	829	2 B27211	virA protein - Agr
80	31	75.6	884	2 T40690	hypothetical prote
81	31	75.6	966	2 T30017	hypothetical prote
82	31	75.6	1006	2 AD2195	hypothetical prote
83	31	75.6	1036	1 GNLJG2	HIV-1 retropepsin
84	31	75.6	1132	2 D88250	protein T21B10.3 [
85	31	75.6	1170	2 T25043	hypothetical prote
86	31	75.6	1213	2 T37959	hypothetical prote
87	31	75.6	1213	2 S45323	genome polyprotein
88	31	75.6	1210	1 JQ1948	genome polyprotein
89	31	75.6	2412	1 JQ1537	genome polyprotein
90	30	73.2	66	2 T12922	hypothetical prote
91	30	73.2	97	2 G34964	ig heavy chain V-I
92	30	73.2	104	2 S26467	ig heavy chain V r
93	30	73.2	115	2 F25114	ig heavy chain V r
94	30	73.2	116	1 HVM331	ig heavy chain pre
95	30	73.2	120	2 A25114	ig heavy chain V r
96	30	73.2	123	2 A99504	hypothetical prote
97	30	73.2	140	2 AB2692	transcription regu
98	30	73.2	140	2 F97473	heavy metal depend
99	30	73.2	174	2 T39995	probable splice fa
100	30	73.2	202	2 G64565	hypothetical prote
101	30	73.2	270	2 T48205	hypothetical prote
102	30	73.2	278	1 A47404	alpha-tocopherol t

103	30	73.2	278	2	S54352	alpha-tocopherol t
104	30	73.2	288	2	S68197	regulatory protein
105	30	73.2	297	2	T38372	RNA binding protein
106	30	73.2	344	2	F86267	T6J4.6 protein - A
107	30	73.2	351	2	JC4752	ferrochelatase (EC
108	30	73.2	355	2	T30873	dNDP-glucose dehyd
109	30	73.2	391	2	AG2536	hypothetical prote
110	30	73.2	396	2	C49904	hypothetical prote
111	30	73.2	396	2	C91189	L-lactate dehydrog
112	30	73.2	396	2	D86036	L-lactate dehydrog
113	30	73.2	437	2	S70429	hypothetical prote
114	30	73.2	499	2	T32337	hypothetical prote
115	30	73.2	516	2	JE0301	inulinase [EC 3.2.
116	30	73.2	545	2	T51438	hypothetical prote
117	30	73.2	549	2	B45154	SMF2 protein - yea
118	30	73.2	570	2	I45712	Dpp receptor SAX p
119	30	73.2	570	2	T16373	hypothetical prote
120	30	73.2	661	2	A69252	3-hydroxyacyl-CoA
121	30	73.2	684	1	RNLVCL	DNA-directed RNA p
122	30	73.2	713	2	JC5870	poly(beta-D-mannur
123	30	73.2	741	2	T13677	NADH2 dehydrogenas
124	30	73.2	984	2	A29513	mineralocorticoid
125	30	73.2	1004	2	S51133	transposase Tn4652
126	30	73.2	1153	2	A97179	ATP-dependent exon
127	30	73.2	1235	1	VG1HMJ	E2 glycoprotein pr
128	30	73.2	1366	2	B86292	F7H2.12 protein -
129	30	73.2	1376	1	VG1HJ2	E2 glycoprotein pr
130	30	73.2	1376	1	JQ1534	E2 glycoprotein pr
131	30	73.2	1677	2	T46095	hypothetical prote
132	30	73.2	2254	2	D86215	protein T6D22.14 [
133	30	73.2	2701	2	T17796	inositol-trisphosp
134	30	73.2	4063	2	T42993	probable spectrin
135	30	73.2	4101	2	T23630	hypothetical prote
136	29	70.7	13	2	PH0805	T-cell receptor al
137	29	70.7	103	2	E81030	hypothetical prote
138	29	70.7	120	2	T10814	lipid transfer pro
139	29	70.7	120	2	T18015	hypothetical prote
140	29	70.7	147	2	S21648	T-cell receptor al
141	29	70.7	152	2	C81262	probable lipoprote
142	29	70.7	162	2	T40662	peripheral-type be
143	29	70.7	183	2	A86819	peptide methionine
144	29	70.7	184	2	B97329	multimeric flavodo
145	29	70.7	185	2	AE1331	hypothetical prote
146	29	70.7	185	2	AE1702	hypothetical prote
147	29	70.7	186	2	F83352	NADH-dependent FMN
148	29	70.7	191	2	B70042	hypothetical prote
149	29	70.7	195	2	C95974	hypothetical prote
150	29	70.7	197	2	E72743	hypothetical prote

ALIGNMENTS

RESULT 1	
I49635	mouse Dhml protein - mouse
C:Species: Mus musculus (house mouse)	
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Nov-1999	
C:Accession: I49635	
R:Shobuike, T.; Sugano, S.; Yamashita, T.; Ikeda, H.	
Nucleic Acids Res. 23, 357-361, 1995	
A:Title: Characterization of cDNA encoding mouse homolog of fission yeast dhpl+ gene: st	
A:Reference number: I49635; MUID:95192042; PMID:7885830	
A:Accession: I49635	
A:Status: preliminary; translated from GB/EMBL/DBJ	
A:Molecule type: mRNA	
A:Residues: 1-947 <RES>	
A:Cross-references: GB:D38517; NID:g1060920; PIDN:BAA07524.1; PID:g1060921	
Query Match	90.2%; Score 37; DB 2; Length 947;
Best Local Similarity	100.0%; Pred. No. 23;
Matches	6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	GNNYRP 6	
Db	843	GNNYRP 848	
RESULT 2			
H64422			
type II restriction enzyme homolog - Methanococcus jannaschii			
C:Species: Methanococcus jannaschii			
C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000			
C:Accession: H64422			
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, A.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.; rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.			
Science 273, 1058-1073, 1996			
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.			
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii			
A:Reference number: A64300; MUID:96337999; PMID:8688087			
A:Accession: H64422			
A:Status: preliminary; nucleic acid sequence not shown; translation not shown			
A:Molecule type: DNA			
A:Residues: 1-222 <BUL>			
A:Cross-references: GB:U67541; GB:L77117; NID:g2826353; PIDN:AAB98987.1; PID:g1591646; T			
C:Genetics:			
A:Map position: FOR915789-916457			
A:Start codon: GTG			
C:Superfamily: Methanobacterium mthZIM protein			
Query Match 85.4%; Score 35; DB 2; Length 222;			
Best Local Similarity 85.7%; Pred. No. 14;			
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
Qy	1	GNNYRPS 7	
Db	147	GNNYRPS 153	
RESULT 3			
S61435			
metalloproteinase (EC 3.4.24.-) precursor - Aspergillus fumigatus			
C:Species: Aspergillus fumigatus			
C>Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 29-Jan-1999			
C:Accession: S61435; S61436; S61450			
R:Jaton-Ogay, K.; Paris, S.; Huerre, M.; Quadroni, M.; Falchetto, R.; Togni, G.; Latge, I.			
Mol. Microbiol. 14, 917-928, 1994			
A:Title: Cloning and disruption of the gene encoding an extracellular metalloprotease of			
A:Reference number: S61435; MUID:95231298; PMID:7715453			
A:Accession: S61435			
A:Molecule type: DNA			
A:Residues: 1-633 <JAT>			
A:Cross-references: EMBL:Z30424; NID:g458475; PID:g458476			
A:Experimental source: strain delta18			
A:Accession: S61436			
A:Molecule type: mRNA			
A:Residues: 1-633 <JAW>			
A:Cross-references: EMBL:Z30424			
A:Accession: S61450			
A:Molecule type: protein			
A:Residues: 246-276; 362-371; 407-416; 419-428; 477-488; 571-588 <JAF>			
C:Keywords: hydrolase; metalloproteinase; zinc; zymogen			
F1-18/Domain: signal sequence #status predicted <SIG>			
F19-245/Domain: propeptide #status predicted <PRO>			
F1246-633/Product: metalloproteinase #status experimental <MAT>			
F1428-432/Binding site: zinc, catalytic (His) (active) #status predicted			
F1429/Active site: Glu #status predicted			
Query Match 85.4%; Score 35; DB 2; Length 633;			
Best Local Similarity 100.0%; Pred. No. 39;			
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	2	NNYRPS 7	
Db	308	NNYRPS 313	

```

RESULT 4
S42894
Metalloproteinase MEP - Aspergillus fumigatus
C:Species: Aspergillus fumigatus
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Sep-1997
C:Accession: S42894
R:Jaton-Ogay, K.; Quadroni, M.; Falchetto, R.; Togni, G.; Paris, S.; Latge, J.; Monod, M.
submitted to the EMBL Data Library, March 1994
A:Description: Nucleotide sequence of genomic and cDNA clones encoding an extracellular
A:Reference number: S42894
A:Accession: S42894
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-649 <JAT>
A:Cross-references: EMBL:Z30424; NID:G458475; PID:G458476
C:Genetics:
A:Introns: 119/3; 251/2; 588/2

Query Match      85.4%; Score 35; DB 2; Length 649;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 NNYRPS 7
      |||||
DB      308 NNYRPS 313

RESULT 5
G71607
probable integral membrane protein PFB0735c - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C:Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 21-Jul-2000
C:Accession: G71607
R:Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.;
; Pertea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.
Science 282, 1126-1132, 1998
A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
A:Reference number: A71600; MUID:99021743; PMID:9804551
A:Accession: G71607
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1182 <GAR>
A:Cross-references: GB:AE001414; GB:AE001362; NID:G3845260; PIDN:RAC71936.1; PID:G3845260
A:Experimental source: Clone 3D7
C:Genetics:
A:Gene: PFB0735c

Query Match      85.4%; Score 35; DB 2; Length 1182;
Best Local Similarity 100.0%; Pred. No. 73;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 NNYRPS 7
      |||||
DB      624 NNYRPS 629

RESULT 6
S45889
probable regulatory protein YBR033w - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein YBR0318
C:Species: Saccharomyces cerevisiae
C:Date: 26-Aug-1994 #sequence_revision 09-Sep-1994 #text_change 19-Apr-2002
R:Grivell, L.A.; de Haan, M.; Maat, M.J.; Smits, P.H.M.
submitted to the Protein Sequence Database, August 1994
A:Reference number: S45875
A:Accession: S45889
A:Molecule type: DNA
A:Residues: 1-919 <GRI>
A:Cross-references: EMBL:Z35902; NID:G536247; PIDN:CAA84975.1; PID:G536248; MIPS:YBR033w
A:Experimental source: strain S288C

```

```

R:Smits, P.H.M.; de Haan, M.; Maat, C.; Grivell, L.A.
Yeast 10(Suppl.A), S75-S80, 1994
A:Title: The complete sequence of a 33 kb fragment on the right arm of chromosome II from
ly identified genes and a homologue of the SCO1 gene.
A:Reference number: S46551; MUID:94378725; PMID:8091864
A:Accession: S46562
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-919 <SMI>
A:Cross-references: EMBL:X76078; NID:G498748; PIDN:CAA53688.1; PID:G498760
A:Experimental source: strain S288C
A:Note: The nucleotide sequence was submitted to the EMBL Data Library, November 1993
C:Genetics:
A:Cross-references: SGD:S0000237
A:Map position: 2R
A:Note: YBR033w
C:Superfamily: unassigned GAL4-type zinc cluster proteins; GAL4 zinc binuclear cluster h
C:Keywords: DNA binding; transmembrane protein; zinc finger
F:51-90/Domain: GAL4 zinc binuclear cluster homology <GAL4>
F:56-85/Region: zinc finger
F:408-424/Domain: transmembrane #status predicted <TM1>
F:482-504/Domain: transmembrane #status predicted <TM2>
F:547-563/Domain: transmembrane #status predicted <TM3>
F:703-720/Domain: transmembrane #status predicted <TM4>
F:790-806/Domain: transmembrane #status predicted <TM5>

Query Match      82.9%; Score 34; DB 2; Length 919;
Best Local Similarity 85.7%; Pred. No. 90;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GNNYRPS 7
      |||||
DB      130 GNNYRPS 136

RESULT 7
S23533
hypothetical protein 311 - diatom (Cylindrotheca fusiformis) plasmid pCF1
C:Species: Cylindrotheca fusiformis
C:Date: 03-May-1994 #sequence_revision 20-Feb-1995 #text_change 21-Jul-2000
R:Hildebrand, M.; Hasegawa, P.; Ord, R.W.; Thorpe, V.S.; Glass, C.A.; Volcani, B.E.
Plant Mol. Biol. 19, 759-770, 1992
A:Title: Nucleotide sequence of diatom plasmids: identification of open reading frames w
A:Reference number: S23531; MUID:92353385; PMID:1322740
A:Accession: S23533
A:Molecule type: DNA
A:Residues: 1-311 <HIL>
A:Cross-references: EMBL:X64302; NID:G17988; PIDN:CRA45582.1; PID:G17991
C:Genetics:
A:Genome: plasmid
C:Superfamily: Cylindrotheca fusiformis plasmid pCF1 hypothetical protein 311

Query Match      80.5%; Score 33; DB 2; Length 311;
Best Local Similarity 83.3%; Pred. No. 49;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 GNNYRP 6
      |||||
DB      108 GNNYQP 113

RESULT 8
T00244
sopB protein - Escherichia coli plasmids p0157 and F
N:Alternate names: plasmid partitioning protein; protein B
C:Species: Escherichia coli
C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 21-Jul-2000
C:Accession: T00244; S28096; T42169
R:Makino, K.; Ishii, K.; Yasunaga, T.; Hattori, M.; Yokoyama, K.; Yatsudo, H.C.; Kubota,
S.; Shinagawa, H.
DNA Res. 5, 1-9, 1998
A:Title: Complete nucleotide sequences of 93-kb and 3.3-kb plasmids of an enterohemorrhag

```

C:Accession: F82843
R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen

A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.

A:Reference number: A82515; MUID:20365717; PMID:10910347

A:Note: For a complete list of authors see reference number A59328 below

A:Accession: F82843

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-491 <SIM>

A:Cross-references: GB:AE003867; GB:AE003849; NID:g9104919; PIDN:AAF82951.1; GSPDB:GN0012

A:Experimental source: strain 9a5c

R:Simpson, A.J.G.; Reinach, P.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; AJ

Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H.

as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.

submitted to GenBank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohne

chado, M.A.; Madeira, A.M.B.N.; Kemp, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuranae, E.E.; Laigre

, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Marques, M.V.; Martins, E.

Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasaka

M.; Tsuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z

A:Reference number: A59328

A:Contents: annotation

C:Genetics:

A:Gene: sopB

A:Genome: plasmid

C:Superfamily: parB protein

C:Keywords: DNA binding; plasmid partition

Query Match 80.5%; Score 33; DB 2; Length 323;

Best Local Similarity 71.4%; Pred. No. 51;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNYRPS 7

Db 152 GNDYRPT 158

RESULT 9

T13113

probable sopB protein - phase N15

N:Alternate names: protein gp27

C:Species: phase N15

C>Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 20-Sep-1999

R:Henarix, R.W.; Ravin, V.K.; Casjens, S.R.; Ford, M.E.; Ravin, N.V.; Smirnov, I.K.

submitted to the EMBL Data Library, May 1998

A:Reference number: Z17603

A:Accession: T13113

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-342 <HEN>

A:Cross-references: EMBL:AF074613; PIDN:AAC70137.1

C:Genetics:

A:Note: Gene 27

C:Superfamily: parB protein

Query Match 80.5%; Score 33; DB 2; Length 342;

Best Local Similarity 71.4%; Pred. No. 54;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNYRPS 7

Db 155 GNDYRPT 161

RESULT 10

F82843

aminopeptidase A/I XF0138 [imported] - Xylella fastidiosa (strain 9a5c)

C:Species: Xylella fastidiosa

C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 03-Mar-2003

C:Accession: F82843

R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen

A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.

A:Reference number: A82515; MUID:20365717; PMID:10910347

A:Note: For a complete list of authors see reference number A59328 below

A:Accession: F82843

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-491 <SIM>

A:Cross-references: GB:AE003867; GB:AE003849; NID:g9104919; PIDN:AAF82951.1; GSPDB:GN0012

A:Experimental source: strain 9a5c

R:Simpson, A.J.G.; Reinach, P.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; AJ

Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H.

as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.

submitted to GenBank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohne

chado, M.A.; Madeira, A.M.B.N.; Kemp, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuranae, E.E.; Laigre

, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Marques, M.V.; Martins, E.

Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasaka

M.; Tsuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z

A:Reference number: A59328

A:Contents: annotation

C:Genetics:

A:Gene: sopB

A:Genome: plasmid

C:Superfamily: parB protein

C:Keywords: DNA binding; plasmid partition

Query Match 80.5%; Score 33; DB 2; Length 323;

Best Local Similarity 71.4%; Pred. No. 51;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNYRPS 7

Db 152 GNDYRPT 158

RESULT 11

I51556

recombination activating protein - African clawed frog

C:Species: Xenopus laevis (African clawed frog)

C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 02-Sep-2000

R:Greenhalgh, P.; Olesen, C.E.; Steiner, L.A.

J. Immunol. 151, 3100-3110, 1993

A:Title: Characterization and expression of recombination activating genes (RAG-1 and RAG

A:Reference number: I51555; MUID:93389137; PMID:8376769

A:Accession: I51556

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-520 <GRE>

A:Cross-references: GB:I19325; NID:g401812; PIDN:AAA49943.1; PID:g401813

C:Genetics:

A:Gene: RAG-2

C:Superfamily: recombination-activating protein RAG-2

Query Match 80.5%; Score 33; DB 2; Length 520;

Best Local Similarity 83.3%; Pred. No. 82;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNYRPS 6

Db 225 GNNFRP 230

RESULT 12

F86875

peptide methionine sulfoxide reductase [imported] - Lactococcus lactis subsp. lactis (str

C:Species: Lactococcus lactis subsp. lactis

C>Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001

C:Accession: F86875

R. Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissbach, J.; Ehrlich, G. *Genome Res.* 11, 731-753, 2001
 A:Title: The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* sensu stricto
 A:Reference number: A86625; MUID:21235186; PMID:11337471
 A:Accession: F86875
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-172 <STO>
 A:Cross-references: GB:AE005176; PID:gl2725052; PIDN:AAK06104.1; GSPDB:GN00146
 A:Experimental source: strain IL1403
 C:Genetics:
 A:Gene: pmsR
 C:Superfamily: peptide methionine sulfoxide reductase

Query Match 78.0%; Score 32; DB 2; Length 172;
 Best Local Similarity 83.3%; Pred. No. 44;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNYRP 6
 DB 92 GDNYP 97
 |||||
 -:-:-

RESULT 13
 A35009
 C 3.4.23.42 thermopain () precursor - *Sulfolobus acidocaldarius*
 C:Species: *Sulfolobus acidocaldarius*
 C:Date: 20-Jul-1990 #sequence_revision 20-Jul-1990 #text_change 03-Jun-2002
 C:Accession: A35009
 R:Lin, X.; Tang, J.
 J. Biol. Chem. 265, 1490-1495, 1990
 A:Title: Purification, characterization, and gene cloning of thermopain, a thermostable
 A:Reference number: A35009; MUID:90110209; PMID:2104843
 A:Accession: A35009
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-340 <LIN>
 A:Cross-references: GB:J05184; NID:gl52940; PID:gl52941
 C:Superfamily: *Sulfolobus acidocaldarius* thermopain
 C:Keywords: hydrolase

Query Match 78.0%; Score 32; DB 2; Length 340;
 Best Local Similarity 71.4%; Pred. No. 85;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
 DB 246 GNNYTPN 252
 |||||
 -:-:-

RESULT 14
 B46216
 Transcription factor Skn-1, splice form i - rat
 N:Alternate names: Skn-1i
 C:Species: *Rattus norvegicus* (Norway rat)
 C:Date: 19-Nov-1993 #sequence_revision 18-Nov-1994 #text_change 05-Feb-1999
 R:Andersen, B.; Schonemann, M.D.; Flynn, S.E.; Pearse II, R.V.; Singh, H.; Rosenfeld, M.
 Science 260, 78-82, 1993
 A:Title: Skn-1a and Skn-1i: two functionally distinct Oct-2-related factors expressed in
 A:Reference number: A46216; MUID:93219836; PMID:7682011
 A:Accession: B46216
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-348 <AND>
 A:Cross-references: GB:L23863; NID:G9393221
 A:Note: sequence extracted from NCBI backbone (NCBIP:128574)
 C:Superfamily: transcription factor Oct-2; homeobox homology; POU domain homology
 C:Keywords: alternative splicing; DNA binding; homeobox; nucleus; transcription regulation
 F:101-168/Domain: POU domain homology <POU>
 F:193-249/Domain: homeobox homology <HOX>

Query Match 78.0%; Score 32; DB 1; Length 348;
 Best Local Similarity 85.7%; Pred. No. 12;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
 DB 379 GNNSRPS 385
 |||||
 -:-:-

RESULT 17
 A46216
 Transcription factor Skn-1, splice form a - rat
 N:Alternate names: Skn-1a
 C:Species: *Rattus norvegicus* (Norway rat)
 C:Date: 19-Nov-1993 #sequence_revision 18-Nov-1994 #text_change 05-Dec-1997
 C:Accession: A46216
 R:Andersen, B.; Schonemann, M.D.; Flynn, S.E.; Pearse II, R.V.; Singh, H.; Rosenfeld, M.

Best Local Similarity 85.7%; Pred. No. 87;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
 DB 297 GNNSRPS 303
 |||||
 -:-:-

RESULT 15
 T49791
 hypothetical protein B9J10.280 [imported] - *Neurospora crassa*
 C:Species: *Neurospora crassa*
 C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
 C:Accession: T49791
 R:Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura, R.;
 submitted to the Protein Sequence Database, May 2000
 A:Reference number: Z25022
 A:Accession: T49791
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-349 <SCH>
 A:Cross-references: EMBL:AL356324; GSPDB:GN00116; NCSP:B9J10.280
 A:Experimental source: BAC clone B9J10; strain OR74A
 C:Genetics:
 A:Gene: NCSP:B9J10.280
 A:Map position: 6

Query Match 78.0%; Score 32; DB 2; Length 349;
 Best Local Similarity 83.3%; Pred. No. 88;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 NNNYRPS 7
 DB 145 NNYRPA 150
 |||||
 -:-:-

RESULT 16
 S35541
 Transcription factor Oct-11, short splice form - mouse
 C:Species: *Mus musculus* (house mouse)
 C:Date: 09-Dec-1993 #sequence_revision 01-Sep-1995 #text_change 22-Jun-1999
 C:Accession: S35541
 R:Goldsbrough, A.S.; Healy, L.E.; Copeland, N.G.; Gilbert, D.J.; Jenkins, N.A.; Willisor
 Nucleic Acids Res. 21, 127-134, 1993
 A:Title: Cloning, chromosomal localization and expression pattern of the POU domain gene
 A:Reference number: S35541; MUID:93181164; PMID:8441607
 A:Accession: S35541
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-403 <GOL>
 A:Cross-references: EMBL:Z18537; NID:G53514; PIDN:CAA79222.1; PID:G53515
 C:Superfamily: transcription factor Oct-2; homeobox homology; POU domain homology
 C:Keywords: alternative splicing; DNA binding; homeobox; nucleus; transcription regulation
 F:275-331/Domain: homeobox homology <HOX>

Query Match 78.0%; Score 32; DB 1; Length 403;
 Best Local Similarity 85.7%; Pred. No. 12+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
 DB 379 GNNSRPS 385
 |||||
 -:-:-

RESULT 17
 A46216
 Transcription factor Skn-1, splice form a - rat
 N:Alternate names: Skn-1a
 C:Species: *Rattus norvegicus* (Norway rat)
 C:Date: 19-Nov-1993 #sequence_revision 18-Nov-1994 #text_change 05-Dec-1997
 C:Accession: A46216
 R:Andersen, B.; Schonemann, M.D.; Flynn, S.E.; Pearse II, R.V.; Singh, H.; Rosenfeld, M.

Science 260, 78-82, 1993
A:Title: Skn-1a and Skn-1i: two functionally distinct Oct-2-related factors expressed in
A:Reference number: A46216; MUID:93219836; PMID:7682011
A:Accession: A46216
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-430 <AND>
A:Cross-references: GB:123862; NID:g393220
A:Note: sequence extracted from NCBI backbone (NCBIP:128572)
C:Superfamily: transcription factor Oct-2; homeobox homology; POU domain homology
C:Keywords: alternative splicing; DNA binding; homeobox; nucleus; transcription regulati
F:183-250/Domain: POU domain homology <POU>
F:275-331/Domain: homeobox homology <HOX>

Query Match 78.0%; Score 32; DB 1; Length 430;
Best Local Similarity 85.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GNNYRPS 7
Db 379 GNNSRPS 385

RESULT 18
JC2002 transcription factor Oct-11, long splice form - mouse
N:Alternate names: Epcoc-1 protein; POU domain protein Oct-11; POU-domain transcription f
C:Species: Mus musculus (house mouse)
C>Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 22-Jun-1999
C:Accession: JC2002; S09239
R:Yukawa, K.; Yasui, T.; Yamamoto, A.; Shiku, H.; Kishimoto, T.; Kikutani, H.
Gene 133, 163-169, 1993
A:Title: Epcoc-1: A POU-domain gene expressed in murine epidermal basal cells and thymic
A:Reference number: JC2002; MUID:9404806; PMID:8224904
A:Accession: JC2002
A:Molecule type: mRNA
A:Residues: 1-431 <YUK>
A:Cross-references: GB:114677; NID:g388913; PIDN:AAA16855.1; PID:g388914
R:Goldsbrough, A.; Ashworth, A.; Wallison, K.
Nucleic Acids Res. 18, 1634, 1990
A:Title: Cloning and sequencing of POU-boxes expressed in mouse testis.
A:Reference number: S09237; MUID:90221898; PMID:1970171
A:Accession: S09239
A:Molecule type: DNA
A:Residues: 200-318 <GOL>
A:Cross-references: EMBL:X51960; NID:g53464; PIDN:CAA36219.1; PID:g9930192
C:Comment: This protein is a developmental regulator that controls organ development and
C:Genetics:
A:Gene: Oct-11
C:Superfamily: transcription factor Oct-2; homeobox homology; POU domain homology
C:Keywords: alternative splicing; DNA binding; homeobox; nucleus; transcription regulati
F:183-250/Domain: POU domain homology <POU>
F:275-331/Domain: homeobox homology <HOX>

Query Match 78.0%; Score 32; DB 1; Length 431;
Best Local Similarity 85.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GNNYRPS 7
Db 379 GNNSRPS 385

RESULT 19
H84862 hypothetical protein At2g43170 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: H84862
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: H84862
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-504 <STO>
A:Cross-references: GB:AE002093; NID:g3763926; PIDN:AAC64306.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g43170
A:Map position: 2

Query Match 78.0%; Score 32; DB 2; Length 504;
Best Local Similarity 83.3%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GNNYRPS 6
Db 481 GNNYRPS 486

RESULT 20
F96491 hypothetical protein T4i21.13 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: F96491
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luroe, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I.
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: F96491
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-530 <STO>
A:Cross-references: GB:AE005173; NID:g10092525; PIDN:AAG12923.1; GSPDB:GN00141
C:Genetics:
A:Gene: T4i21.13
A:Map position: 1

Query Match 78.0%; Score 32; DB 2; Length 530;
Best Local Similarity 83.3%; Pred. No. 1.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 2 GNNYRPS 7
Db 55 GNNYRPS 60

RESULT 21
F90564 conserved hypothetical protein MYPV_4220 [imported] - Mycoplasma pulmonis (strain UAB CT)
C:Species: Mycoplasma pulmonis
C>Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 03-Aug-2001
C:Accession: F90564
R:Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer, I.;
Nucleic Acids Res. 29, 2145-2153, 2001
A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulm
A:Reference number: A99512; MUID:21267165; PMID:11353084
A:Accession: F90564
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-572 <KUR>
A:Cross-references: GB:AL445366; PID:g14089836; PIDN:CAC13595.1; GSPDB:GN00153
A:Experimental source: strain UAB CTIP
C:Genetics:
A:Gene: MYPV_4220

A:Genetic code: SGC3

Query Match 78.0%; Score 32; DB 2; Length 572;
 Best Local Similarity 83.3%; Pred. No. 1.4e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 GNNYRPS 7
 |||||
 Db 33 GNNYRPT 38

RESULT 22

AB2362

ribonuclease II [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp. PCC 7120

A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002

C:Accession: AB2362

R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriiguchi, N.; Nakazaki, N.; Shimpō, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S. DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AB2362

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-686 <KUR>

A:Cross-references: GB:BA0000019; PIDN:BA076149.1; PID:g17133586; GSPDB:GN00179

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: all4450

Query Match 78.0%; Score 32; DB 2; Length 686;
 Best Local Similarity 71.4%; Pred. No. 1.7e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
 |||||
 Db 52 GNTYKPS 58

RESULT 23

AD1857

hypothetical protein allo405 [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp. PCC 7120

A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002

C:Accession: AD1857

R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriiguchi, N.; Nakazaki, N.; Shimpō, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S. DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AD1857

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-692 <KUR>

A:Cross-references: GB:BA0000019; PIDN:BA072363.1; PID:g17129750; GSPDB:GN00179

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: all0405

Query Match 78.0%; Score 32; DB 2; Length 692;
 Best Local Similarity 83.3%; Pred. No. 1.7e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNYRP 6
 |||||
 Db 373 GNDYRP 378

RESULT 24

T52564

Mre11 protein homolog [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 24-Oct-2000 #sequence_revision 24-Oct-2000 #text_change 15-Sep-2003

C:Accession: T52564

R:Hartung, F.

submitted to the EMBL Data Library, July 1999

A:Reference number: Z26116

A:Accession: T52564

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-720 <HAR>

A:Cross-references: EMBL:AJ243822; PIDN:CAB50793.1

C:Genetics:

A:Note: MRE11

C:Superfamily: double-strand break repair protein MRE11; phosphoesterase core homology

Query Match 78.0%; Score 32; DB 2; Length 720;
 Best Local Similarity 71.4%; Pred. No. 1.8e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
 |||||
 Db 296 GNYRPT 302

RESULT 25

E84475

probable Athila retroelement ORF1 protein [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C:Accession: E84475

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;

Euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: AB4420; MUID:20083487; PMID:10617197

A:Accession: E84475

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-750 <STO>

A:Cross-references: GB:AE002093; NID:g4388721; PIDN:AAD19759.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2G06160

A:Map position: 2

Query Match 78.0%; Score 32; DB 2; Length 750;
 Best Local Similarity 83.3%; Pred. No. 1.9e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 NNYRPS 7
 |||||
 Db 295 NNYRPN 300

Search completed: September 24, 2004, 01:54:28

Job time : 15.434 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2005, 17:57:45 ; Search time 52.7188 Seconds
(without alignments)
51.354 Million cell updates/sec

Title: US-10-088-639a-2_COPY_49_55
Perfect score: 41
Sequence: 1 GNNYRPS 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 385760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04: *
1: Geneseqp1980s: *
2: Geneseqp1990s: *
3: Geneseqp2000s: *
4: Geneseqp2001s: *
5: Geneseqp2002s: *
6: Geneseqp2003as: *
7: Geneseqp2003bs: *
8: Geneseqp2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	100.0	7	ADM09121	Canine im
2	41	100.0	249	AAB68087	An anti-a
3	37	90.2	182	AB864583	Drosophil
4	37	90.2	182	ADL15285	Fruitfly
5	37	90.2	865	AB863795	AB863795 Drosophil
6	37	90.2	930	ABU53195	Human nuc
7	36	87.8	7	AAB40011	Anti-hi1
8	36	87.8	201	AAB16608	Human nov
9	36	87.8	201	ABU55677	Human nov
10	36	87.8	249	ADH44202	sfv antib
11	36	87.8	438	ABG10678	Novel hum
12	36	87.8	585	AAY06237	Human tra
13	36	87.8	588	ABG92868	Clock pro
14	36	87.8	591	AAU16186	Human nov
15	36	87.8	591	ABU55255	Human nov
16	36	87.8	599	ABG92867	Clock pro
17	36	87.8	602	AAU79162	Aryl hydr
18	36	87.8	602	ABU56486	Lung canc
19	36	87.8	602	ADK52095	Human ato
20	36	87.8	602	ADL14099	Human sar
21	36	87.8	602	ADN04976	Antipsori
22	36	87.8	622	ABG92866	Clock pro
23	36	87.8	636	ABG92879	BNAL2 (br
24	36	87.8	636	ABG92865	Clock pro
25	35	85.4	108	ADI22141	Anti-plat

26	35	85.4	249	8	ADI58094	Reg IV-ep
27	35	85.4	251	5	ABP45544	Human Bly
28	35	85.4	251	7	ADG96371	Single ch
29	35	85.4	1182	3	AAB18288	Plasmodiu
30	34	82.9	919	5	ABP35687	Fungal 2B
31	34	82.9	919	8	ADN18782	Bacterial
32	33	80.5	7	8	ADQ91421	CDR2 of t
33	33	80.5	110	4	AAG80220	Human aut
34	33	80.5	112	8	ADQ91398	Amino aci
35	33	80.5	155	4	ABG21683	Novel hum
36	33	80.5	157	4	ABG15752	Novel hum
37	33	80.5	157	7	ABO64892	Klebsiell
38	33	80.5	158	4	ABG21677	Novel hum
39	33	80.5	214	4	ABG24566	Novel hum
40	33	80.5	220	4	ABG21701	Novel hum
41	33	80.5	221	4	ABG24565	Novel hum
42	33	80.5	227	4	ABG10676	Novel hum
43	33	80.5	235	4	ABG04022	Novel hum
44	33	80.5	240	4	ABG24588	Novel hum
45	33	80.5	243	6	AAO31145	Human CMO

ALIGNMENTS

RESULT 1
ADM09121
ID ADM09121 standard; peptide; 7 AA.
XX

AC ADM09121;

DT 20-MAY-2004 (first entry)

XX Canine immunoglobulin light chain variable domain CDR2 peptide 58.

XX canine; dog; heavy; immunoglobulin; antibody light chain variable domain;
KW antiallergic; allergy; IgE; gene therapy;
KW complementarity determining region; CDR2.

XX Canis familiaris.

XX WO2003060080-A2.

XX 24-JUL-2003.

XX 20-DEC-2002; 2002WO-US041362.

XX 21-DEC-2001; 2001US-0344874P.

XX (INDEX-) IDEXX LAB INC.

XX Krah ER, Guo H, Aiyappa A, Lawton R;

XX WPI; 2003-598521/56.

XX New canine heavy and light chain variable domain polypeptides, useful for
treating canine allergy.

XX Disclosure; Page 24; 130pp; English.

XX The invention relates to a novel canine heavy or light chain variable
domain polypeptide. The protein of the invention demonstrates
antiallergic activity and may be useful for treating canine allergy,
possibly via gene therapy. The current sequence is that of a canine
immunoglobulin light chain variable domain complementarity determining
region (CDR) peptide of the invention.

XX Sequence 7 AA;

Query Match 100.0%; Score 41; DB 7; Length 7;
Best Local Similarity 100.0%; Pred. NO. 1.8e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
 DB 1 GNNYRPS 7

RESULT 2
 AAB68087
 ID AAB68087 standard; protein; 249 AA.
 XX AAB68087;
 AC AAB68087;
 XX 09-JUL-2001 (first entry)
 DE An anti-alpha6beta4 integrin light chain linked to a heavy chain.
 XX Gastrointestinal epithelial tumour cell; alpha6beta4 integrin;
 KW tumour-associated antigen; metastatic disease; malignant disease;
 KW tumour typing; tumour screening; tumour.
 XX Synthetic.
 OS Macaca fascicularis.
 XX Key Location/Qualifiers
 FT Region 23..33
 FT /note= "Complementarity determining region (CDR) 1 of the
 FT light chain"
 FT 45..55
 FT /note= "Complementarity determining region (CDR) 2 of the
 FT light chain"
 FT 88..98
 FT /note= "Complementarity determining region (CDR) 3 of the
 FT light chain"
 FT 110..127
 FT /note= "linker"
 FT 158..162
 FT /note= "Complementarity determining region (CDR) 1 of the
 FT heavy chain"
 FT 177..193
 FT /note= "Complementarity determining region (CDR) 1 of the
 FT heavy chain"
 FT 226..238
 FT /note= "Complementarity determining region (CDR) 1 of the
 FT heavy chain"
 XX WO200130854-A2.
 XX 03-MAY-2001.
 XX 26-OCT-2000; 2000WO-SE002082.
 XX 28-OCT-1999; 99SE-00003895.
 XX (ACTI-) ACTIVE BIOTECH AB.
 XX Brodin TN, Karlstroem PJ, Ohlsson LG, Tordsson MJ, Kearney PP;
 PI Nilson BHK;
 XX WPI; 2001-308619/32.
 DR N-PSDB; AAF84797.
 XX Novel antibody for diagnosis, treatment of human metastatic and malignant
 PT diseases, has binding structure for target structure displayed on cell
 PT surface of human gastrointestinal epithelial tumor and normal cells.
 XX Claim 1; Page 55-56; 75pp; English.
 XX The present sequence represents a Monkey antibody light chain linked to a
 CC heavy chain. The antibody binds to a target structure displayed in and on
 CC the cell surface of human gastrointestinal epithelial tumour cells and in
 CC a subpopulation of normal human gastrointestinal epithelial cells. The
 CC target structure especially comprises alpha6beta4 integrin. This is a
 CC tumour-associated antigen. The antibody, and its fragments, are useful
 CC for treating conditions based on anti-angiogenic mechanism in humans.

CC They are useful for treating human metastatic and malignant disease, for
 CC in vitro, in vivo diagnosis and prognosis of human malignant disease,
 CC comprising tumour typing, tumour screening, tumour diagnosis and
 CC prognosis and monitoring premalignant conditions. Quantitative in vivo
 CC diagnosis is carried out by determining the localization of antibody to
 CC tumour deposits in humans
 XX Sequence 249 AA;
 SQ Query Match 100.0%; Score 41; DB 4; Length 249;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GNNYRPS 7
 DB 49 GNNYRPS 55
 RESULT 3
 ABB64583
 ID ABB64583 standard; protein; 182 AA.
 XX ABB64583;
 AC ABB64583;
 XX 26-MAR-2002 (first entry)
 DT Drosophila melanogaster polypeptide SEQ ID NO 20541.
 XX Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX Drosophila melanogaster.
 OS WO200171042-A2.
 XX 27-SEP-2001.
 PD 23-MAR-2001; 2001WO-US009231.
 PF 23-MAR-2000; 2000US-0191637P.
 PR 11-JUL-2000; 2000US-00614150.
 XX (PEKE) PE CORP NY.
 PA Venter JC, Adams M, Li PWD, Myers EW;
 PI WPI; 2001-656860/75.
 DR N-PSDB; ABL08686.
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions.
 XX Disclosure; SEQ ID NO 20541; 21pp + Sequence Listing; English.
 XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 182 AA;
 SQ Query Match 90.2%; Score 37; DB 4; Length 182;
 Best Local Similarity 100.0%; Pred. No. 80;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GNNYRPS 6

```

Db      134 GNNYRP 139
|||||
RESULT 4
ADD15285
ID ADD15285 standard; protein; 182 AA.
XX AC
XX ADD15285;
XX DT
XX 15-JAN-2004 (first entry)
XX DE
XX Fruitfly odourant receptor protein (SeqId 37).
XX KW
XX odourant receptor; fruitfly; mating; repulsion; flight; insect damage;
XX KW disease spread; pesticide; insect management program.
XX OS
XX Drosophila melanogaster.
XX PN
XX WO2003020913-A2.
XX PD
XX 13-MAR-2003.
XX PF
XX 04-SEP-2002; 2002WO-US028315.
XX PR
XX 04-SEP-2001; 2001US-0317401P.
XX PA (SENT-) SENTIGEN CORP.
XX PI
XX Lee KJ, Ong J, Nguyen TT, Kloss B;
XX DR
XX WPI; 2003-300885/29.
XX PT
XX Novel Anopheles gambiae odourant receptor polypeptides and nucleic acid
XX PT encoding the polypeptides, useful as targets for identifying pest control
XX PT agents.
XX PS
XX Claim 1; SEQ ID NO 37; 172pp; English.
XX CC
XX This invention relates to novel Anopheles gambiae odourant receptor genes
XX CC and encoded proteins thereof. Specifically, it refers to the isolated
XX CC genes of the African malaria mosquito that are related to the 'classical'
XX CC Drosophila odourant receptor genes, and compounds that bind to and
XX CC modulate these receptors cause various behavioural responses such as
XX CC mating, repulsion or flight. Accordingly, the present invention describes
XX CC such compounds (both natural and synthetic) that are useful for
XX CC attracting insects to traps or to localised toxins, for repelling insects
XX CC from individuals or populated residential areas, or for interfering with
XX CC the function of olfactory system such that insects are unable to locate
XX CC food and hosts. As such, these compounds can be used to control insect
XX CC damage and the spread of disease, and will significantly reduce
XX CC dependence on toxic pesticides having a direct and immediate impact on
XX CC coordinated insect management programs. This polypeptide sequence is a
XX CC Drosophila melanogaster odourant receptor protein, which is homologous to
XX CC the African malaria mosquito proteins of the invention.
XX SQ
XX Sequence 182 AA;

Query Match 90.2%; Score 37; DB 7; Length 182;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNYRP 6
Db 134 GNNYRP 139
|||||
RESULT 5
ABB63795
ID ABB63795 standard; protein; 865 AA.
XX AC
XX ABB63795;
XX XX

26-MAR-2002 (first entry)
XX DE
XX Drosophila melanogaster polypeptide SEQ ID NO 18177.
XX KW
XX Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical.
XX OS
XX Drosophila melanogaster.
XX PN
XX WO200171042-A2.
XX PD
XX 27-SEP-2001.
XX PF
XX 23-MAR-2001; 2001WO-US009231.
XX PR
XX 23-MAR-2000; 2000US-0191637P.
XX PR
XX 11-JUL-2000; 2000US-00614150.
XX XX
XX (PEKE ) PE CORP NY.
XX PA
XX Venter JC, Adams M, Li PWD, Myers EW;
XX PI
XX WPI; 2001-656860/75.
XX DR
XX N-PSDB; ABL07898.
XX XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX PT genes from Drosophila and for elucidating cell signaling and cell-cell
XX PT interactions.
XX PS
XX Disclosure; SEQ ID NO 18177; 21pp + Sequence Listing; English.
XX CC
XX The invention relates to an isolated nucleic acid detection reagent
XX CC capable of detecting 1000 or more genes from Drosophila. The invention is
XX CC useful in developmental biology and in elucidating cell signalling and
XX CC cell-cell interactions in higher eukaryotes for the development of
XX CC insecticides, therapeutics and pharmaceutical drugs. The invention
XX CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
XX CC ABB72072). The sequence data for this patent did not form part of the
XX CC printed specification, but was obtained in electronic format directly
XX CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ
XX Sequence 865 AA;

Query Match 90.2%; Score 37; DB 4; Length 865;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNYRP 6
Db 424 GNNYRP 429
|||||
RESULT 6
ABU53195
ID ABU53195 standard; protein; 930 AA.
XX AC
XX ABU53195;
XX XX
XX 14-APR-2003 (first entry)
XX DE
XX Human nucleic acid management-associated DKFZphtes3_2ml8 homologue.
XX KW
XX Human; gene therapy; vaccine; disease treatment; detection.
XX OS
XX Homo sapiens.
XX PN
XX WO200112659-A2.
XX XX
XX 22-FEB-2001.
XX PD
XX 18-AUG-2000; 2000WO-IB001496.
XX PF
XX

```

PR 18-AUG-1999; 99US-0149499P.
XX 28-SEP-1999; 99US-0156503P.
PA (GEHU-) GERMAN HUMAN GENOME PROJECT.
XX
PI Wiemann S;
XX
XX WPI; 2001-327840/34.
XX
XX Nucleic acids having the sequences of clones isolated from libraries of
PT different human tissues, useful in recombinant DNA methodologies.
XX
XX Example III; Page 809; 1095pp; English.
XX
XX This invention describes novel polynucleotides and polypeptides isolated
CC from human cDNA libraries which can be used for gene therapy or in
CC vaccines. The polynucleotides of the invention and antibodies encoded by
CC them may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate polypeptide expression. The products of the
CC invention may also be used to identify modulators of expression and
CC activity and to down regulate expression and activity. The antibodies of
CC the invention may also be used as diagnostic agents for detecting the
CC presence of polypeptides in samples. This sequence represents a homologue
CC of a polypeptide described in the disclosure of the invention
XX
SQ Sequence 930 AA;

Query Match 90.2%; Score 37; DB 4; Length 930;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 GNNYRP 6
DB 843 GNNYRP 848
|||||

RESULT 7
AAB40011
ID AAB40011 standard; peptide; 7 AA.
XX
AC AAB40011;
XX
DT 05-FEB-2001 (first entry)
DE
DE Anti-HIL2 antibody light chain CDR2 amino acid sequence SEQ ID 527.
XX
XX Human; neutralising antibody; interleukin-12; IL-12; antiinflammatory;
KW complementarity determining region; CDR; antirheumatic; antiarthritic;
KW antisclerotic; neuroprotective; antipsoriatic; antiasthmatic; cardiant;
KW antiparasitic; antibacterial; immunosuppressive; Crohn's disease;
KW multiple sclerosis; rheumatoid arthritis.
XX
OS Homo sapiens.
XX
XX WO200056772-A1.
XX
XX 28-SEP-2000.
XX
XX 24-MAR-2000; 2000WO-US007946.
XX
XX 25-MAR-1999; 99US-0126603P.
XX
XX (BADI) BASF AG.
PA (GEMY) GENETICS INST INC.
XX
XX Salfeld JG, Roguska M, Paskind M, Banerjee S, Tracey DE, White M;
PI Kaymakcan Z, Labkovsky B, Sakorafas P, Friedrich S, Myles A;
PI Veldman GM, Venturini A, Warne NW, Widom A, Elvin JG, Duncan AR;
PI Derbyshire EJ, Carmen S, Smith S, Holtet TL, Du Fou SL;
XX
XX WPI; 2000-638250/61.
XX
XX New human antibody specific for human interleukin-12 (IL-12) used to

PT	treat disorders characterized by aberrant IL-12 expression e.g. Crohn's
PT	disease and multiple sclerosis.
XX	
PS	Claim 33; Fig 2G; 377pp; English.
XX	
CC	This invention relates to a new human antibody specific for human
CC	interleukin-12 (IL-12). The invention also includes antigen binding
CC	portions that bind to IL-12. Sequences AAB39485-BJ9516 represent human
CC	anti-IL-12 antibody heavy and light chain complementarity determining
CC	region (CDR) amino acid sequences, and also includes variable region
CC	amino acid sequences. Other variable region amino acid sequences are
CC	given in AAB39517-B39560 and AAB40088-B40149. Sequences AAB39561-B39771
CC	represent anti-IL-12 CDR3 related amino acid sequences, AAB39772-B40063
CC	represent other CDR sequences. Light chain CDR3 consensus sequences are
CC	given in AAB40064-B40067. Primers used in the identification and
CC	construction of the antibodies of the invention are given in AAC61062-
CC	C61071. The antibody of the invention is a neutralising antibody and has
CC	antirheumatic; antiarthritic; antisclerotic; antiinflammatory;
CC	neuroprotective; antiporiatic; antitasthmatic; cardiant; antiparasitic;
CC	antibacterial and immunosuppressive activity. The antibodies or antigen-
CC	binding fragments are useful in the treatment of disorders associated
CC	with detrimental release of human IL-12, especially Crohn's disease,
CC	multiple sclerosis and rheumatoid arthritis. They can also be used in the
CC	manufacture of a pharmaceutical composition to treat human IL-12
CC	disorders
XX	
SQ	Sequence 7 AA;
	Query March 87.8%; Score 36; DB 3; Length 7;
	Best Local Similarity 85.7%; Pred. No. 1.8e+06;
	Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0
QY	1 GNYYRPS 7
DB	1 GNDYRPS 7
RESULT 8	
AUU16608	
ID	AAU16608 standard; protein; 201 AA.
XX	
AC	AAU16608;
XX	
DT	07-NOV-2001 (first entry)
XX	
DE	Human novel secreted protein, Seq ID 1561.
XX	
KW	Human; immunosuppressive; antiarthritic; antirheumatic; cytostatic;
KW	cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective;
KW	antibacterial; virucide; fungicide; opthalmological; vulnerary;
KW	secreted protein; rheumatoid arthritis; hyperproliferative disorder;
KW	cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW	cerebral ischaemia; angiogenesis; nervous system disorder;
KW	Alzheimer's disease; infection; ocular disorder; corneal infection;
KW	wound healing; epithelial cell proliferation; skin ageing; food additive;
KW	preservative; antiproliferative.
XX	
OS	Homo sapiens.
XX	
FN	WO200155322-A2.
XX	
PD	02-AUG-2001.
XX	
PF	17-JAN-2001; 2001WO-US001341.
XX	
FR	31-JAN-2000; 2000US-0179065P.
PR	04-FEB-2000; 2000US-0180628P.
PR	24-FEB-2000; 2000US-0184664P.
PR	02-MAR-2000; 2000US-0186350P.
PR	16-MAR-2000; 2000US-0189874P.
PR	17-MAR-2000; 2000US-0190076P.
PR	18-APR-2000; 2000US-0198123P.
PR	19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225113P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225477P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 23-AUG-2000; 2000US-0227182P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0232403P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.

PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241828P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-488783/53.

N-PSDB; AAS26595.

New nucleic acid molecules encoding 461 human secreted proteins for
diagnosing, preventing, treating or ameliorating medical conditions and
used as food additives or preservatives.

Claim 11; SEQ ID NO 1561; 980pp; English.

The invention relates to isolated nucleic acid molecules and their
encoded secreted proteins. The nucleic acids and proteins are used to
prevent, treat or ameliorate a medical condition in e.g. humans, mice,
rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used

CC in diagnosing a pathological condition or susceptibility to a
 CC pathological condition. Antibodies to the proteins can also be used in
 CC alleviating symptoms associated with the disorders and in diagnostic
 CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays
 CC (ELISA). Disorders which are diagnosed or treated include autoimmune
 CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
 CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
 CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,
 CC nervous system disorders e.g. Alzheimer's disease, infections caused by
 CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection,
 CC and many other disorders listed in the specification. The polypeptides
 CC can also be used to aid wound healing and epithelial cell proliferation,
 CC to prevent skin aging due to sunburn, to maintain organs before
 CC transplantation, for supporting cell culture of primary tissues, to
 CC regenerate tissues and in chemotaxis. The polypeptides can also be used
 CC as a food additive or preservative to increase or decrease storage
 CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
 CC minerals, cofactors and other nutritional components. The present
 CC sequence represents a novel secreted protein of the invention. Note: The
 CC sequence data for this patent did not form part of the printed

Query Match 87.8%; Score 36; DB 4; Length 201;

Best Local Similarity 85.7%; Pred. No. 1.3e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNYRPS 7

DB 124 GSNYRPS 130

RESULT 9

ABU55677 ID ABU55677 standard; protein; 201 AA.

XX AC ABU55677;

DT 18-MAR-2003 (first entry)

DE Human novel polypeptide #764.

XX Human; neural disorder; immune system disorder; renal disorder;
 KW muscular disorder; respiratory disease; reproductive disorder;
 KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
 KW hyperproliferative disorder; inflammatory disease; allergic reaction;
 KW blood related disorder; cancer; immunosuppressive; antiinflammatory;
 KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;
 KW haemostatic; antiarteriosclerotic.

OS Homo sapiens.

XX PN US2002132753-A1.

XX PD 19-SEP-2002.

XX PF 17-JAN-2001; 2001US-00764864.

XX PR 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 28-JUN-2000; 2000US-0214886P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225757P.

PR 14-AUG-2000; 2000US-0225758P.
 PR 22-AUG-2000; 2000US-0226868P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 13-OCT-2000; 2000US-0237040P.
 PR 20-OCT-2000; 2000US-0239935P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.

(ROSE/) ROSEN C A.

(RUBE/) RUBEN S M.

(BARA/) BARASH S C.

Rosen CA, Ruben SM, Barash SC;

WPI; 2003-147444/14.

N-PSDB; ABX73936.

XX New polypeptides and nucleic acids, useful in gene therapy for treating,
 PT inhibiting or preventing e.g. neural, immune system, muscular,
 PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or
 PT renal disorders.
 XX
 PS Claim 11; SEQ ID NO 1561; 402pp; English.

XX The invention relates to human novel polypeptides and their associated

CC polynucleotides. The polypeptides and polynucleotides are useful in gene
 CC therapy for treating, inhibiting or preventing neural disorders, immune
 CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis
 CC and multiple sclerosis), muscular disorders, respiratory diseases (e.g.
 CC nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,
 CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders,
 CC (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left
 CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage
 CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and
 CC leukaemia), inflammatory diseases (e.g. septic shock, bursitis and
 CC appendicitis), allergic reactions and conditions (e.g. asthma), blood
 CC related disorders (e.g. thrombosis, atherosclerosis and myocardial
 CC infarction) and cancerous diseases. Sequences ABU54914-ABU55699 and
 CC ABU55748 represent human novel polypeptides of the invention

XX Sequence 201 AA;

Query Match 87.8%; Score 36; DB 6; Length 201;
 Best Local Similarity 85.7%; Pred. No. 1.3e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNYRPS 7

Db 124 GSNYRPS 130
|:|||||
RESULT 10
ADH44202
ID ADH44202 standard; protein; 249 AA.
AC ADH44202;
XX
XX
DT 25-MAR-2004 (first entry)
XX
DE sFv antibody fragment from clone 42.
XX
XX antibody; single-chain fragment; human; replication inhibitor;
KW hepatitis C virus; variable region; VL; VH; light chain; heavy chain;
KW viral protein; E1; E2; core protein; NS3-protease; NS3-helicase;
KW NS4A cofactor; NS5B RNA polymerase; virucide; hepatotropic;
KW antiinflammatory; vaccine; passive immunisation; diagnosis; gene therapy.
XX
OS Homo sapiens.
FH Key Location/Qualifiers
FT Region 7..28
FT /note= "FR-L1"
FT Region 29..39
FT /note= "CDR-L1"
FT Region 40..54
FT /note= "FR-L2"
FT Region 55..61
FT /note= "CDR-L2"
FT Region 62..93
FT /note= "FR-L3"
FT Region 94..103
FT /note= "CDR-L3"
FT Region 104..116
FT /note= "J-segment"
FT Region 117..126
FT /note= "Linker fragment"
FT Region 127..156
FT /note= "FR-H1"
FT Region 157..162
FT /note= "CDR-H1"
FT Region 163..176
FT /note= "FR-H2"
FT Region 177..192
FT /note= "CDR-H2"
FT Region 193..224
FT /note= "FR-H3"
FT Region 225..249
FT /note= "CDR-H3"
XX
XX WO200293519-A2.
PN
XX
XX 21-NOV-2002.
XX
XX 13-MAY-2002; 2002WO-EP005227.
PF
XX
PR 11-MAY-2001; 2001DE-01023041.
XX
XX (HEIN/) HEINTGES T.
XX
XX Heintges T, Tessmann K, Hauselinger D, Artsaenko O;
XX WPI; 2003-120726/11.
DR
XX
XX New single-chain human antibody fragment, useful for treating or
PT diagnosing hepatitis C virus infection, has affinity for an essential
PT viral protein.
XX
XX Claim 8; SEQ ID NO 58; 78pp; German.
PS
XX
XX This invention describes a novel single-chain fragment of human antibody

CC that inhibits replication of hepatitis C virus (HCV) and comprises the
CC variable regions (VL and VH) of the light and heavy chains of an antibody
CC directed against at least one essential viral protein. The invention also
CC describes a method of identifying antibody fragments that inhibit
CC replication of HCV. The protein is an envelope, core or non-structural
CC protein, especially E1, E2, core, NS3-protease or -helicase, NS4A
CC cofactor or NS5B RNA polymerase (especially a NS3 helicase). The VL and
CC VH regions may be linked covalently or through a linker, particularly a
CC hydrophilic and/or flexible peptide. The single chain fragment may also
CC be linked to a non-protein polymer and where intended for diagnosis, it
CC may include disulfide bridges. The method of the invention involves
CC identifying a DNA library of antibody fragments against one or more
CC essential HCV proteins prepared from bone marrow of patients with chronic
CC HCV infection. The library is cloned and expressed on the surface of
CC phages. Phages that express fragments with high affinity for essential
CC proteins are selected by many cycles of selection and reamplification
CC (panning). Sequences that encode the selected fragments are then
CC expressed, as soluble proteins, in bacteria (particularly Escherichia
CC coli) and recovered from the periplasmic space. The products of the
CC invention have virucide, hepatotropic and antiinflammatory activity. The
CC novel single chain fragments are used to prepare vaccines, especially for
CC passive immunisation and for diagnosis of HCV infection or for gene. The
CC DNA encoding the single chain fragment is useful for gene therapy of HCV.
CC The sequence and structure of single chain fragments can be used for
CC design of HCV protein inhibitors. This sequence represents a human
CC antibody fragment described in the disclosure of the invention.
XX
XX Sequence 249 AA;
SQ

Query Match 87.8%; Score 36; DB 7; Length 249;
Best Local Similarity 85.7%; Pred. No. 1.7e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GNNYRPS 7
DB 55 GNNHRPS 61
|||:||||
RESULT 11
ABG10678
ID ABG10678 standard; protein; 438 AA.
XX
XX AC ABG10678;
XX
XX DT 13-FEB-2002 (first entry)
XX
XX DE Novel human diagnostic protein #10669.
XX
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX OS Homo sapiens.
XX
XX PN WO200175067-A2.
XX
XX PD 11-OCT-2001.
XX
XX PF 30-MAR-2001; 2001WO-US008631.
XX
XX PR 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.
XX
XX PA (HYSE-) HYSEQ INC.
XX
XX PI Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
DR N-PSDB; AAS74865.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.

XX PS Claim 20; SEQ ID NO 41037; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, CC and in recombinant production of (II). The polynucleotides are also used CC in diagnostics as expressed sequence tags for identifying expressed CC genes. (I) is useful in gene therapy techniques to restore normal CC activity of (II) or to treat disease states involving (II). (II) is CC polypeptide in tissue, as molecular weight markers and as a food CC supplement. (II) and its binding partners are useful in medical imaging CC of sites expressing (II). (I) and (II) are useful for treating disorders CC involving aberrant protein expression or biological activity. The CC polypeptide and polynucleotide sequences have applications in CC diagnostics, forensics, gene mapping, identification of mutations CC and to produce other types of data and products dependent on DNA and CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic CC amino acid sequences. The invention. Note: The sequence data for this CC patent did not appear in the printed specification, but was obtained in CC electronic format directly from WIPO at CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 438 AA;

Query Match 87.8%; Score 36; DB 4; Length 438;
Best Local Similarity 85.7%; Pred. No. 2.9e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
DB 174 GNDYRPS 180
|:|||||

RESULT 12
AA06297
ID AAY06297 standard; protein; 585 AA.
AC AAY06297;
XX
XX 23-AUG-1999 (first entry)
DE Human transcription regulator MOP9.
XX MOP9; member of the PAS superfamily; bHLH-PAS; human;
KW transcription regulator; hypoxia.
XX Homo sapiens.
OS
XX WO928464-A2.
XX
XX 10-JUN-1999.
XX
XX 27-NOV-1998; 98WO-US025314.
XX
XX 28-NOV-1997; 97US-0066863P.
XX (WISC) WISCONSIN ALUMNI RES FOUND.
XX
XX Bradfield CA, Gu YZ, Hogenesch JB;
XX
XX WPI; 1999-371120/31.
XX N-PSDB; AAX58988.
XX Developmental signal transduction associated proteins.
PT
XX Claim 5; Page 106; 106pp; English.
XX
XX The present sequence represents MOP9, a novel member of the PAS
XX superfamily, where PAS stands for PER/ARNT/SIM domains. MOP9 cDNA (see
XX AAX58988) was cloned from human brain mRNA. MOP9 displays extended

CC homology to MOP3 (see AAY06291). It pairs with CLOCK and MOP4 (see
CC AAY06292) and binds an E-box element with flanking region specificity.
CC The invention provides novel MOP 2-9 nucleic acids (see AAX58981-88) and
CC proteins (see AAY06289-97). These are useful in a variety of research,
CC diagnostic and therapeutic applications. Several of the MOPs are alpha-
CC class hypoxia-inducible factors. Others are involved in circadian signal
CC transduction
XX
SQ Sequence 585 AA;

Query Match 87.8%; Score 36; DB 2; Length 585;
Best Local Similarity 85.7%; Pred. No. 3.8e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
DB 118 GSNYRPS 124
|:|||||

RESULT 13
ABG92868
ID ABG92868 standard; protein; 588 AA.
XX
XX ABG92868;
XX
XX 29-AUG-2003 (revised)
DT 19-NOV-2002 (first entry)
XX
XX Clock protein BAML2 (brain-muscle-Arnt-like protein 2) #4.
XX Human; clock protein BMAL2; brain-muscle-Arnt-like protein 2; insomnia;
KW sleeping disorder; non-24-hour sleep; sleep-phase forward;
KW retreat syndrome; time-zone variation syndrome.
XX
XX Homo sapiens.
XX
XX WO200264785-A1.
XX
XX 22-AUG-2002.
XX
XX 23-AUG-2001; 2001WO-JP007197.
XX
XX 13-FEB-2001; 2001JP-00035743.
XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX
XX Fukada Y, Okano T;
XX
XX WPI; 2002-667007/71.
XX N-PSDB; ABS68504.
XX
XX Clock gene Baml2 and expressed clock protein BMAL2 important in clock
XX oscillation mechanism and relating to circadian rhythm, used in diagnosis
XX of and developing drugs for insomnia and other sleeping disorders.
XX
XX Claim 1; Page 102-105; 187pp; Japanese.
XX
XX The invention relates to a DNA sequence encoding clock protein BMAL2
XX (brain-muscle-Arnt-like protein 2). The gene and protein are applicable
XX in diagnosis of and development of drugs for insomnia and other sleeping
XX disorders e.g. non-24-hour sleep, sleep-phase forward or retreat syndrome
XX and time-zone variation syndrome. ABG92865-ABG92879 represent BMAL2 amino
XX acid sequences of the invention. (Updated on 29-AUG-2003 to standardise
XX OS field)
XX
SQ Sequence 588 AA;

Query Match 87.8%; Score 36; DB 5; Length 588;
Best Local Similarity 85.7%; Pred. No. 3.8e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
|:|||||

Db 121 GSNRPS 127

RESULT 14
AAU16186
ID AAU16186 standard; protein; 591 AA.
AC XX
AC AAU16186;
DT DT
DT 07-NOV-2001 (first entry)
XX DE
XX Human novel secreted protein, Seq ID 1139.
XX Human; immunosuppressive; antiarthritic; antirheumatic; cytostatic;
KW cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective;
KW antibacterial; virucide; fungicide; ophthalmologic; vulnerary;
KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW cerebral ischaemia; angiogenesis; nervous system disorder;
KW Alzheimer's disease; infection; ocular disorder; corneal infection;
KW wound healing; epithelial cell proliferation; skin ageing; food additive;
KW preservative; antiproliferative.
OS Homo sapiens.
XX
PN WO200155322-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001341.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225475P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.

17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251388P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX
 (HUMA-) HUMAN GENOME SCI INC.
 PA
 Rosen CA, Barash SC, Ruben SM;
 PI
 WPI; 2001-488783/53.
 DR N-PSDB; AAS26173.
 DR
 XX
 New nucleic acid molecules encoding 461 human secreted proteins for
 PT diagnosing, preventing, treating or ameliorating medical conditions and
 PT used as food additives or preservatives.
 PT
 XX
 Claim 11; SEQ ID NO 1139; 980pp; English.
 PS
 XX
 The invention relates to isolated nucleic acid molecules and their
 CC encoded secreted proteins. The nucleic acids and proteins are used to
 CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
 CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
 CC in diagnosing a pathological condition or susceptibility to a
 CC pathological condition. Antibodies to the proteins can also be used in
 CC alleviating symptoms associated with the disorders and in diagnostic
 CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays
 CC (ELISA). Disorders which are diagnosed or treated include autoimmune
 CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
 CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
 CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,
 CC nervous system disorders e.g. Alzheimer's disease, infections caused by
 CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection,
 CC and many other disorders listed in the specification. The polypeptides
 CC can also be used to aid wound healing and epithelial cell proliferation,
 CC to prevent skin aging due to sunburn, to maintain organs before
 CC transplantation, for supporting cell culture of primary tissues, to
 CC regenerate tissues and in chemotaxis. The polypeptides can also be used
 CC as a food additive or preservative to increase or decrease storage
 CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
 CC minerals, cofactors and other nutritional components. The present
 CC sequence represents a novel secreted protein of the invention. Note: The
 CC sequence data for this patent did not form part of the printed

Query Match 87.8%; Score 36; DB 4; Length 591;
 Best Local Similarity 85.7%; Pred. NO. 3.9e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GNNYRPS 7
 Db 124 GSNYRPS 130

RESULT 15
 ABUS5255
 ID ABUS5255 standard; protein; 591 AA.
 XX

AC
 XX ABUS5255;
 DT 18-MAR-2003 (first entry)
 XX
 DE Human novel polypeptide #342.
 XX
 KW Human; neural disorder; immune system disorder; renal disorder;
 KW muscular disorder; respiratory disease; reproductive disorder;
 KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
 KW hyperproliferative disorder; inflammatory disease; allergic reaction;
 KW blood related disorder; cancer; immunosuppressive; antiinflammatory;
 KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;
 KW haemostatic; antiarteriosclerotic.
 XX
 OS Homo sapiens.
 XX
 PN US2002132753-A1.
 XX
 PD 19-SEP-2002.
 XX
 PF 17-JAN-2001; 2001US-00764864.
 XX
 PR 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 07-JUL-2000; 2000US-0218880P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 11-JUL-2000; 2000US-0217496P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 22-AUG-2000; 2000US-0226868P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 08-SEP-2000; 2000US-0229513P.
 PR 21-SEP-2000; 2000US-0231413P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239935P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 XX

PA (ROSE/) ROSEN C A.
 PA (RUBE/) RUBEN S M.
 PA (BARA/) BARASH S C.
 XX
 PI Rosen CA, Ruben SM, Barash SC;
 XX
 DR WPI; 2003-147444/14.
 DR N-PSDB; ABX73514.
 XX
 PT New polypeptides and nucleic acids, useful in gene therapy for treating,
 PT inhibiting or preventing e.g. neural, immune system, muscular,
 PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or
 PT renal disorders.
 XX
 PS Claim 11; SEQ ID NO 1139; 402pp; English.
 XX
 CC The invention relates to human novel polypeptides and their associated
 CC polynucleotides. The polypeptides and polynucleotides are useful in gene
 CC therapy for treating, inhibiting or preventing neural disorders, immune
 CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis
 CC and multiple sclerosis), muscular disorders, respiratory diseases (e.g.
 CC nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,
 CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
 CC (e.g. congenital heart defects, Epstein's anomaly and hypoplastic left
 CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage
 CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and
 CC leukaemia), inflammatory diseases (e.g. septic shock, bursitis and
 CC appendicitis), allergic reactions and conditions (e.g. asthma), blood
 CC related disorders (e.g. thrombosis, atherosclerosis and myocardial
 CC infarction) and cancerous diseases. Sequences ABU54914-ABU55699 and
 CC ABU55748 represent human novel polypeptides of the invention
 XX
 SQ Sequence 591 AA;

Query Match 87.8%; Score 36; DB 6; Length 591;
 Best Local Similarity 85.7%; Pred. No. 3.8e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GNNYRPS 7
 Db 124 GSNYRPS 130
 |.|||||

Search completed: April 28, 2005, 18:17:21
 Job time : 55.7188 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2005, 18:06:50 ; Search time 14.2188 Seconds
(without alignments)
36.750 Million cell updates/sec

Title: US-10-088-639A-2_COPY_49_55
Perfect score: 41
Sequence: 1 GNNYRPS 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/prodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/prodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/prodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/prodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/prodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37	90.2	157	4	US-09-270-767-44201
2	36	87.8	602	3	US-09-374-454-19
3	33	80.5	157	4	US-09-489-039A-11409
4	33	80.5	324	4	US-09-489-039A-11252
5	32	78.0	132	2	US-08-345-321-4
6	32	78.0	159	4	US-09-621-976-4449
7	32	78.0	223	4	US-09-107-532A-7214
8	32	78.0	348	1	US-08-035-392-2
9	32	78.0	348	1	US-08-504-511A-2
10	32	78.0	430	1	US-08-035-392-4
11	32	78.0	430	1	US-08-504-511A-4
12	32	78.0	442	4	US-09-540-236-3486
13	32	78.0	552	4	US-09-835-654-2
14	32	78.0	612	4	US-09-248-796A-24557
15	32	78.0	720	3	US-09-480-921B-8
16	32	78.0	1115	4	US-09-976-594-893
17	31	75.6	7	2	US-08-672-345C-38
18	31	75.6	7	2	US-08-672-345C-74
19	31	75.6	7	3	US-09-214-095D-38
20	31	75.6	7	3	US-09-214-095D-74
21	31	75.6	74	4	US-09-621-976-4090
22	31	75.6	75	4	US-09-621-976-5643
23	31	75.6	105	4	US-09-538-092-153
24	31	75.6	109	2	US-08-672-345C-3
25	31	75.6	109	2	US-08-672-345C-93
26	31	75.6	109	3	US-09-214-095D-3
27	31	75.6	109	3	US-09-214-095D-121

28	31	75.6	134	4	US-09-248-796A-19693	Sequence 19693, A
29	31	75.6	175	4	US-09-270-767-46152	Sequence 46152, A
30	31	75.6	202	4	US-09-107-433-4846	Sequence 4846, Ap
31	31	75.6	330	4	US-09-489-039A-10928	Sequence 10928, A
32	31	75.6	388	2	US-08-894-772-2	Sequence 2, Appli
33	31	75.6	388	2	US-09-207-844-2	Sequence 2, Appli
34	31	75.6	396	4	US-09-107-532A-5033	Sequence 5033, Ap
35	31	75.6	398	4	US-09-583-110-3408	Sequence 3408, Ap
36	31	75.6	433	4	US-09-809-665A-161	Sequence 161, App
37	31	75.6	517	4	US-09-902-540-11116	Sequence 1116, A
38	31	75.6	632	1	US-08-443-104-4	Sequence 4, Appli
39	31	75.6	632	1	US-08-442-859-4	Sequence 4, Appli
40	31	75.6	632	2	US-08-398-489-4	Sequence 4, Appli
41	31	75.6	632	5	PCT-US95-05534-4	Sequence 4, Appli
42	31	75.6	752	4	US-09-583-110-2714	Sequence 2714, Ap
43	31	75.6	755	4	US-09-107-433-4628	Sequence 4628, Ap
44	31	75.6	1198	3	US-09-199-637A-405	Sequence 405, App
45	30	73.2	39	4	US-09-443-780C-42	Sequence 42, Appli

ALIGNMENTS

RESULT 1
US-09-270-767-44201
; Sequence 44201, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 44201
; LENGTH: 157
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-44201

Query Match 90.2%; Score 37; DB 4; Length 157;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNYR 6
DB 99 GNNYR 104

RESULT 2
US-09-374-454-19
; Sequence 19, Application US/09374454
; Patent No. 6395548
; GENERAL INFORMATION:
; APPLICANT: Lee, Mu-En
; APPLICANT: Maemura, Koji
; APPLICANT: Hsieh, Chung-Ming
; TITLE OF INVENTION: METHODS OF MODULATING OF ANGIOGENESIS
; FILE REFERENCE: 05433/037001
; CURRENT APPLICATION NUMBER: US/09/374,454
; CURRENT FILING DATE: 1999-08-13
; EARLIER APPLICATION NUMBER: US 60/096,515
; EARLIER FILING DATE: 1998-08-14
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 602
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-374-454-19

Query Match 87.8%; Score 36; DB 3; Length 602;

Db 37 GNNRPS 43

RESULT 7

US-09-107-532A-7214
; Sequence 7214, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 7214:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 223 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...223
; SEQUENCE DESCRIPTION: SEQ ID NO: 7214:
US-09-107-532A-7214

Query Match 78.0%; Score 32; DB 4; Length 223;
Best Local Similarity 83.3%; Pred. No. 2.7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNRPS 6
; : : : :
Db 141 GDNYP 146

RESULT 8

US-08-035-392-2
; Sequence 2, Application US/08035392
; Patent No. 5484732
; GENERAL INFORMATION:
; APPLICANT: Rosenfeld, M. G.
; TITLE OF INVENTION: TRANSCRIPTION FACTOR FOR THE REGULATION
; OF THE DEVELOPMENT OF SKIN AND HAIR
; NUMBER OF SEQUENCES: 4

US-08-035-392-2

Query Match 78.0%; Score 32; DB 1; Length 348;
Best Local Similarity 85.7%; Pred. No. 4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNNRPS 7
; : : : :
Db 297 GNNRPS 303

RESULT 9

US-08-504-511A-2
; Sequence 2, Application US/08504511A
; Patent No. 5561224
; GENERAL INFORMATION:
; APPLICANT: Rosenfeld, M. G.
; TITLE OF INVENTION: TRANSCRIPTION FACTOR FOR THE REGULATION
; OF THE DEVELOPMENT OF SKIN AND HAIR
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: California
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/504,511A
; FILING DATE: 20-JUL-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Howells, Stacy L.
; REGISTRATION NUMBER: 34,842
; REFERENCE/DOCKET NUMBER: 07340/002002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 678-5070
; TELEFAX: (619) 678-5099

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 348 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-504-511A-2

Query Match 78.0%; Score 32; DB 1; Length 348;
Best Local Similarity 85.7%; Pred. No. 4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
Db 297 GNNSRPS 303

RESULT 10
US-08-035-392-4
; Sequence 4, Application US/08035392
; Patent No. 5484732
; GENERAL INFORMATION:
; APPLICANT: Rosenfeld, M. G.
; APPLICANT: Andersen, B.
; TITLE OF INVENTION: TRANSCRIPTION FACTOR FOR THE REGULATION
; TITLE OF INVENTION: OF THE DEVELOPMENT OF SKIN AND HAIR
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 1880 Century Park East - Suite 500
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90067

COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: 19930322
APPLICATION NUMBER: US/08/035,392
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Howells, Stacy L.
REGISTRATION NUMBER: 34,842
REFERENCE/DOCKET NUMBER: PD-2607
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 455-5100
TELEFAX: (619) 455-5110
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 430 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-035-392-4

Query Match 78.0%; Score 32; DB 1; Length 430;
Best Local Similarity 85.7%; Pred. No. 4.9e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
Db 379 GNNSRPS 385

RESULT 11
US-08-504-511A-4
; Sequence 4, Application US/08504511A
; Patent No. 5561224
; GENERAL INFORMATION:
; APPLICANT: Rosenfeld, M. G.

; APPLICANT: Andersen, B.
; TITLE OF INVENTION: TRANSCRIPTION FACTOR FOR THE REGULATION
; TITLE OF INVENTION: OF THE DEVELOPMENT OF SKIN AND HAIR
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: California
; COUNTRY: USA
; ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/504,511A
FILING DATE: 20-JUL-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Howells, Stacy L.
REGISTRATION NUMBER: 34,842
REFERENCE/DOCKET NUMBER: 07340/002002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 678-5070
TELEFAX: (619) 678-5099
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 430 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-504-511A-4

Query Match 78.0%; Score 32; DB 1; Length 430;
Best Local Similarity 85.7%; Pred. No. 4.9e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
Db 379 GNNSRPS 385

RESULT 12
US-09-540-236-3486
; Sequence 3486, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRH
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 3486
; LENGTH: 442
; TYPE: PRT
; ORGANISM: M.catarrhialis
US-09-540-236-3486

Query Match 78.0%; Score 32; DB 4; Length 442;
Best Local Similarity 85.7%; Pred. No. 5e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
Db 213 GNNSRPS 219

RESULT 13
US-09-835-654-2

; Sequence 2, Application US/09835654
; Patent No. 6646182
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; TITLE OF INVENTION: Mre11 Orthologue and Uses Thereof
; FILE REFERENCE: 1264
; CURRENT APPLICATION NUMBER: US/09/835,654
; CURRENT FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: US 60/198,570
; PRIOR FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 552
; TYPE: PRT
; ORGANISM: Zea mays
US-09-835-654-2

Query Match 78.0%; Score 32; DB 4; Length 552;
Best Local Similarity 71.4%; Pred. No. 6.1e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
DB 296 GNOYRPT 302

RESULT 14

US-09-248-796A-24557
; Sequence 24557, Application US/09248796A
; Patent No. 6747137

; GENERAL INFORMATION:

; APPLICANT: Keith Weinstock et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN

; FILE REFERENCE: 107196.132

; CURRENT APPLICATION NUMBER: US/09/248,796A

; CURRENT FILING DATE: 1999-02-12

; PRIOR APPLICATION NUMBER: US 60/074,725

; PRIOR FILING DATE: 1998-02-13

; PRIOR APPLICATION NUMBER: US 60/096,409

; PRIOR FILING DATE: 1998-08-13

; NUMBER OF SEQ ID NOS: 28208

; SEQ ID NO 24557

; LENGTH: 612

; TYPE: PRT

; ORGANISM: Candida albicans

; FEATURE:

; NAME/KEY: UNSURE

; LOCATION: (594)

; OTHER INFORMATION: Identity of amino acid sequences at the above locations are unkno

US-09-248-796A-24557

Query Match 78.0%; Score 32; DB 4; Length 612;
Best Local Similarity 71.4%; Pred. No. 6.7e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
DB 455 GKNYKPS 461

RESULT 15

US-09-480-921B-8

; Sequence 8, Application US/09480921B

; Patent No. 6387637

; GENERAL INFORMATION:

; APPLICANT: Levin, Joshua Z.

; APPLICANT: Budziszewski, Gregory J.

; APPLICANT: Potter, Sharon L.

; APPLICANT: Wegrich, Lynette M.

; TITLE OF INVENTION: Herbicide Target Genes and Methods

; FILE REFERENCE: PB/5-30780A

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2005, 18:24:46 ; Search time 43.4219 Seconds
(without alignments)
53.700 Million cell updates/sec

Title: US-10-088-639a-2_COPY_49_55

Perfect score: 41

Sequence: 1 GNNYRPS 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1426032 seqs, 333106140 residues

Total number of hits satisfying chosen parameters: 1426032

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

- Published Applications AA:*
- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
 - 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
 - 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
 - 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
 - 13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
 - 14: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
 - 15: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
 - 16: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
 - 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
 - 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
 - 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
 - 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41	100.0	7	16	US-10-327-598-656
2	36	87.8	7	16	Sequence 527, App
3	36	87.8	201	9	US-10-884-830-527
4	36	87.8	588	15	Sequence 1561, App
5	36	87.8	591	9	US-10-467-721-8
6	36	87.8	599	15	Sequence 1139, App
7	36	87.8	602	14	US-10-467-721-6
8	36	87.8	602	15	Sequence 6, Appli
9	36	87.8	622	15	Sequence 19, Appl
10	36	87.8	636	15	Sequence 78, Appl
11	35	85.4	123	17	US-10-467-721-2
12	35	85.4	251	10	US-10-783-311-334
13	35	85.4	251	15	Sequence 334, App
					Sequence 1555, App
					Sequence 1555, App

14	34	82.9	919	15	US-10-369-493-1435	Sequence 1435, Ap
15	34	82.9	919	15	US-10-149-310-238	Sequence 238, App
16	33	80.5	110	16	US-10-723-434-53	Sequence 53, Appl
17	33	80.5	243	14	US-10-322-673-51	Sequence 51, Appl
18	33	80.5	247	10	US-09-880-748-1993	Sequence 1993, Ap
19	33	80.5	247	10	US-09-880-748-1998	Sequence 1998, Ap
20	33	80.5	247	15	US-10-293-418-1993	Sequence 1993, Ap
21	33	80.5	247	15	US-10-293-418-1998	Sequence 1998, Ap
22	33	80.5	247	16	US-10-779-461-44	Sequence 44, Appl
23	33	80.5	248	10	US-09-880-748-1976	Sequence 1976, Ap
24	33	80.5	248	10	US-09-880-748-2000	Sequence 2000, Ap
25	33	80.5	248	15	US-10-293-418-1976	Sequence 1976, Ap
26	33	80.5	248	15	US-10-293-418-2000	Sequence 2000, Ap
27	33	80.5	250	17	US-10-935-290-117	Sequence 117, App
28	33	80.5	250	17	US-10-935-290-136	Sequence 136, App
29	33	80.5	254	10	US-09-880-748-1958	Sequence 1958, Ap
30	33	80.5	254	15	US-10-293-418-1958	Sequence 1958, Ap
31	33	80.5	258	10	US-09-880-748-2083	Sequence 2083, Ap
32	33	80.5	258	15	US-10-293-418-2083	Sequence 2083, Ap
33	33	80.5	832	15	US-10-094-749-1919	Sequence 1919, Ap
34	33	80.5	861	15	US-10-451-168-75	Sequence 75, Appl
35	33	80.5	922	15	US-10-451-168-77	Sequence 77, Appl
36	33	80.5	1274	9	US-09-746-491-12	Sequence 12, Appl
37	32	78.0	7	14	US-10-203-754A-14	Sequence 14, Appl
38	32	78.0	7	16	US-10-327-598-626	Sequence 626, App
39	32	78.0	7	16	US-10-327-598-698	Sequence 698, App
40	32	78.0	16	17	US-10-735-916A-10	Sequence 10, Appl
41	32	78.0	58	15	US-10-424-539-249492	Sequence 249492,
42	32	78.0	111	17	US-10-727-155-114	Sequence 114, App
43	32	78.0	116	14	US-10-203-754A-20	Sequence 20, Appl
44	32	78.0	117	17	US-10-735-916A-69	Sequence 69, Appl
45	32	78.0	117	17	US-10-735-916A-75	Sequence 75, Appl

ALIGNMENTS

RESULT 1
US-10-327-598-656
; Sequence 656, Application US/10327598
; Publication No. US20040181039A1
; GENERAL INFORMATION:
; APPLICANT: Krah, Eugene
; APPLICANT: Guo, Honliang
; APPLICANT: Aiyappa, Ashok
; APPLICANT: Lawton, Robert
; TITLE OF INVENTION: Canine Immunoglobulin Variable Domains, Caninized Antibodies, and
; TITLE OF INVENTION: for Making and Using Them
; FILE REFERENCE: 01-799-A
; CURRENT APPLICATION NUMBER: US/10/327,598
; CURRENT FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: US 60/344, 874
; PRIOR FILING DATE: 2001-12-21
; NUMBER OF SEQ ID NOS: 1139
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 656
; LENGTH: 7
; TYPE: PRT
; ORGANISM: canis familiaris;
US-10-327-598-656

Query Match 100.0%; Score 41; DB 16; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.3e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
DB 1 GNNYRPS 7

RESULT 2
US-10-884-830-527
; Sequence 527, Application US/10884830

```
; Publication No. US20050004354A1
; GENERAL INFORMATION:
; APPLICANT: Jochen, Salfeld et al.
; TITLE OF INVENTION: Human Antibodies That Bind Human IL-12 And Methods For Producing
; FILE REFERENCE: BBI-093CP
; CURRENT APPLICATION NUMBER: US/10/884,830
; CURRENT FILING DATE: 2004-07-01
; PRIOR APPLICATION NUMBER: US/09/534,717
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/126,603
; PRIOR FILING DATE: March 25, 1999
; NUMBER OF SEQ ID NOS: 675
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 527
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-884-830-527

Query Match      87.8%; Score 36; DB 16; Length 7;
Best Local Similarity 85.7%; Pred. No. 1.3e+06;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 GNNYRPS 7
Db      1 GNDYRPS 7

RESULT 3
US-09-764-864-1561
; Sequence 1561, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1561
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (176)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (195)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1561

Query Match      87.8%; Score 36; DB 9; Length 201;
Best Local Similarity 85.7%; Pred. No. 69;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 GNNYRPS 7
Db      124 GSNYRPS 130

RESULT 4
US-10-467-721-8
; Sequence 8, Application US/10467721
; Publication No. US20040058366A1
; GENERAL INFORMATION:
; APPLICANT: JAPAN SCIENCE AND TECHNOLOGY CORPORATION
; TITLE OF INVENTION: Bmal2, novel clock genes
; FILE REFERENCE: A011-15PCT
; CURRENT APPLICATION NUMBER: US/10/467,721
; CURRENT FILING DATE: 2003-08-11
```

```
; PRIOR APPLICATION NUMBER: JP 2001/35743
; PRIOR FILING DATE: 2001-02-13
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 588
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-467-721-8

Query Match      87.8%; Score 36; DB 15; Length 588;
Best Local Similarity 85.7%; Pred. No. 2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 GNNYRPS 7
Db      121 GSNYRPS 127

RESULT 5
US-09-764-864-1139
; Sequence 1139, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1139
; LENGTH: 591
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (338)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (376)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (465)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (485)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (491)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1139

Query Match      87.8%; Score 36; DB 9; Length 591;
Best Local Similarity 85.7%; Pred. No. 2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 GNNYRPS 7
Db      124 GSNYRPS 130

RESULT 6
US-10-467-721-6
; Sequence 6, Application US/10467721
; Publication No. US20040058366A1
; GENERAL INFORMATION:
; APPLICANT: JAPAN SCIENCE AND TECHNOLOGY CORPORATION
; TITLE OF INVENTION: Bmal2, novel clock genes
; FILE REFERENCE: A011-15PCT
; CURRENT APPLICATION NUMBER: US/10/467,721
; CURRENT FILING DATE: 2003-08-11
```

```

; PRIOR APPLICATION NUMBER: JP 2001/35743
; PRIOR FILING DATE: 2001-02-13
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 599
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-467-721-6

Query Match
Best Local Similarity 87.8%; Score 36; DB 15; Length 599;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
DB 132 GSNYRPS 138

RESULT 7
US-10-121-235-19
; Sequence 19, Application US/10121235
; Publication No. US20030032609A1
; GENERAL INFORMATION:
; APPLICANT: Lee, Mu-En
; APPLICANT: Maemura, Koji
; APPLICANT: Hsieh, Chung-Ming
; TITLE OF INVENTION: METHODS OF MODULATING OF ANGIOGENESIS
; FILE REFERENCE: 05433/037001
; CURRENT APPLICATION NUMBER: US/10/121,235
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: US 09/374,454
; PRIOR FILING DATE: 1999-08-13
; PRIOR APPLICATION NUMBER: US 60/096,515
; PRIOR FILING DATE: 1998-08-14
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 602
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-121-235-19

Query Match
Best Local Similarity 87.8%; Score 36; DB 14; Length 602;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
DB 135 GSNYRPS 141

RESULT 8
US-10-260-708-78
; Sequence 78, Application US/10260708
; Publication No. US20040063101A1
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew
; APPLICANT: Lee, Sang-Yull
; APPLICANT: Old, Lloyd
; TITLE OF INVENTION: Human Sarcoma-Associated Antigens
; FILE REFERENCE: L00461/70138
; CURRENT APPLICATION NUMBER: US/10/260,708
; CURRENT FILING DATE: 2002-09-30
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 78
; LENGTH: 602
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-260-708-78

Query Match
Best Local Similarity 87.8%; Score 36; DB 15; Length 602;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
DB 135 GSNYRPS 141

RESULT 9
US-10-467-721-4
; Sequence 4, Application US/10467721
; Publication No. US20040058366A1
; GENERAL INFORMATION:
; APPLICANT: JAPAN SCIENCE AND TECHNOLOGY CORPORATION
; TITLE OF INVENTION: Bmal2, novel clock genes
; FILE REFERENCE: A011-15PCT
; CURRENT APPLICATION NUMBER: US/10/467,721
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: JP 2001/35743
; PRIOR FILING DATE: 2001-02-13
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 622
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-467-721-4

Query Match
Best Local Similarity 87.8%; Score 36; DB 15; Length 622;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
DB 155 GSNYRPS 161

RESULT 10
US-10-467-721-2
; Sequence 2, Application US/10467721
; Publication No. US20040058366A1
; GENERAL INFORMATION:
; APPLICANT: JAPAN SCIENCE AND TECHNOLOGY CORPORATION
; TITLE OF INVENTION: Bmal2, novel clock genes
; FILE REFERENCE: A011-15PCT
; CURRENT APPLICATION NUMBER: US/10/467,721
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: JP 2001/35743
; PRIOR FILING DATE: 2001-02-13
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 636
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-467-721-2

Query Match
Best Local Similarity 87.8%; Score 36; DB 15; Length 636;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
DB 169 GSNYRPS 175

RESULT 11
US-10-783-311-334
; Sequence 334, Application US/10783311
; Publication No. US20050009136A1
; GENERAL INFORMATION:
; APPLICANT: Nixon, Andrew
; APPLICANT: Hogan, Shannon

```

```

; TITLE OF INVENTION: PAPP-A LIGANDS
; FILE REFERENCE: 10280-059001
; CURRENT APPLICATION NUMBER: US/10/783,311
; CURRENT FILING DATE: 2004-02-19
; PRIOR APPLICATION NUMBER: US 60/448,515
; PRIOR FILING DATE: 2003-02-19
; NUMBER OF SEQ ID NOS: 394
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 334
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Light Chain amino acid sequence
US-10-783-311-334

Query Match      85.4%; Score 35; DB 17; Length 123;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 NNYRPS 7
Db      58 NNYRPS 63
      |||||

RESULT 12
US-09-880-748-1555
; Sequence 1555, Application US/09880748
; Publication No. US20030059937A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523
; CURRENT APPLICATION NUMBER: US/09/880,748
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 3239
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1555
; LENGTH: 251
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-748-1555

Query Match      85.4%; Score 35; DB 10; Length 251;
Best Local Similarity 85.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GKNYRPS 7
Db      191 GKNYRPS 197
      |||||

RESULT 13
US-10-293-418-1555
; Sequence 1555, Application US/10293418
; Publication No. US2003023996A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523P2
; CURRENT APPLICATION NUMBER: US/10/293,418
; CURRENT FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: 60/331,469

```

```

; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/340,817
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 09/880,748
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1555
; LENGTH: 251
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-418-1555

Query Match      85.4%; Score 35; DB 15; Length 251;
Best Local Similarity 85.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GKNYRPS 7
Db      191 GKNYRPS 197
      |||||

RESULT 14
US-10-369-493-1435
; Sequence 1435, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 1435
; LENGTH: 919
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-369-493-1435

Query Match      82.9%; Score 34; DB 15; Length 919;
Best Local Similarity 85.7%; Pred. No. 7.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GKNYRPS 7
Db      130 GKNYRPS 136
      |||||

RESULT 15
US-10-149-310-238
; Sequence 238, Application US/10149310
; Publication No. US20040077039A1
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas
; APPLICANT: Madden, Kevin T.
; APPLICANT: Maxon, Mary
; APPLICANT: Sherman, Amir

```

; TITLE OF INVENTION: Modulation of Secondary Metabolite Production by

; FILE OF INVENTION: Zinc Binuclear Cluster Proteins

; FILE REFERENCE: 14184-019US1

; CURRENT APPLICATION NUMBER: US/10/149,310

; CURRENT FILING DATE: 2003-02-19

; PRIOR APPLICATION NUMBER: PCT/US01/29288

; PRIOR FILING DATE: 2001-09-19

; PRIOR APPLICATION NUMBER: US 60/233,564

; PRIOR FILING DATE: 2000-09-19

; NUMBER OF SEQ ID NOS: 308

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 238

; LENGTH: 919

; TYPE: PRT

; ORGANISM: Saccharomyces cerevisiae

US-10-149-310-238

Query Match 82.9%; Score 34; DB 15; Length 919;
 Best Local Similarity 85.7%; Pred. No. 7.2e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNNYRPS 7

Db 130 GNNILFS 136

Search completed: April 28, 2005, 19:02:13

Job time : 44.4219 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2005, 18:05:55 ; Search time 9.625 Seconds
(without alignments)
69.976 Million cell updates/sec

Title: US-10-088-639A-2_COPY_49_55
Perfect score: 41
Sequence: 1 GNNYRPS 7

Scoring table: BLOSUM62
Gapop 10.0., Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	90.2	947	2 I49635	mouse dhml protein
2	35	85.4	222	2 H64422	type II restrictio
3	35	85.4	633	2 S61435	metalloproteinase
4	35	85.4	649	2 S42894	metalloproteinase
5	35	85.4	1182	2 G71607	probable integral
6	34	82.9	919	2 S45889	probable regulator
7	33	80.5	311	2 S23533	hypothetical prote
8	33	80.5	323	2 T00244	sopB protein - Esc
9	33	80.5	342	2 T13113	probable sopB prot
10	33	80.5	491	2 F82843	aminopeptidase A/I
11	33	80.5	520	2 I51556	recombination acti
12	32	78.0	172	2 F86875	peptide methionine
13	32	78.0	340	2 A35009	C 3.4.23.42 thermo
14	32	78.0	348	1 B42116	transcription fact
15	32	78.0	349	2 T49791	hypothetical prote
16	32	78.0	403	1 S35541	transcription fact
17	32	78.0	430	1 A46216	transcription fact
18	32	78.0	431	1 JC2002	transcription fact
19	32	78.0	504	2 H84862	hypothetical prote
20	32	78.0	530	2 F96491	hypothetical prote
21	32	78.0	572	2 F90564	conserved hypothet
22	32	78.0	686	2 AB2362	ribonuclease II (i
23	32	78.0	692	2 AD1857	hypothetical prote
24	32	78.0	720	2 T52564	Mre11 protein homo
25	32	78.0	750	2 E84475	probable Athila re
26	32	78.0	935	2 S66306	hypothetical prote
27	31	75.6	105	2 S52680	ribosomal protein
28	31	75.6	108	2 T50348	ribosomal protein
29	31	75.6	184	2 S75635	cobN protein - Syn

ALIGNMENTS

RESULT 1

I49635
mouse dhml protein - mouse
C:Species: Mus musculus (house mouse)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Nov-1999
C:Accession: I49635
R:Shobuike, T.; Sugano, S.; Yamashita, T.; Ikeda, H.
Nucleic Acids Res. 23, 357-361, 1995
A:Title: Characterization of cDNA encoding mouse homolog of fission yeast dhpl+ gene: str
A:Reference number: I49635; MUID:95192042; PMID:7885830
A:Accession: I49635
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-947 <RES>
A:Cross-references: GB:D38517; NID:gl060920; PIDN:BAA07524.1; PID:gl060921

Query Match 90.2%; Score 37; DB 2; Length 947;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNYRP 6
DB 843 GNNYRP 848

RESULT 2

H64422
type II restriction enzyme homolog - Methanococcus jannaschii
C:Species: Methanococcus jannaschii
C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
C:Accession: H64422
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, I.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.; rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii.
A:Reference number: A64300; MUID:96337999; PMID:8688087
A:Accession: H64422
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-222 <BUL>
A:Cross-references: UNIPROT:Q58391; GB:U67541; GB:L77117; NID:g2826353; PIDN:AAB98987.1;
C:Genetics:
A:Map position: FOR915789-916457
A:Start codon: GTG
C:Superfamily: Methanobacterium mthZIM protein

Query Match 85.4%; Score 35; DB 2; Length 222;
Best Local Similarity 85.7%; Pred. No. 15;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

carbonate dehydrat
hypothetical prote
probable exported
probable exported
probable transcrip
protein ZC53.6 (lm
carbonate dehydrat
carbonate dehydrat
2-dehydro-3-deoxysg
2-dehydro-3-deoxysg
hypothetical prote
2-dehydro-3-deoxysg
11m protein - Staph
F21H2.9 protein -
uncharacterized pr

QY 1 GNNYRPS 7
 Db 147 GNNYRPS 153

RESULT 3
 S61435
 metalloproteinase (EC 3.4.24.-) precursor - Aspergillus fumigatus
 C:Species: Aspergillus fumigatus
 C>Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 29-Jan-1999
 C:Accession: S61435; S61436; S61450
 R:Jaton-Ogay, K.; Paris, S.; Huere, M.; Quadroni, M.; Falchetto, R.; Togni, G.; Latge, M.
 Mol. Microbiol. 14, 917-928, 1994
 A:Title: Cloning and disruption of the gene encoding an extracellular metalloprotease of
 A:Reference number: S61435; MUID:95231298; PMID:7715453
 A:Accession: S61435
 A:Molecule type: DNA
 A:Residues: 1-633 <JAT>
 A:Cross-references: EMBL:Z30424; NID:G458475; PID:G458476
 A:Experimental source: strain delta18
 A:Accession: S61436
 A:Molecule type: mRNA
 A:Residues: 1-633 <JAW>
 A:Cross-references: EMBL:Z30424
 A:Accession: S61450
 A:Molecule type: protein
 A:Residues: 246-276;362-371;407-416;419-428;477-488;571-588 <JAP>
 C:Keywords: hydrolase; metalloproteinase; zinc; zymogen
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-245/Domain: propeptide #status predicted <PRO>
 F:246-633/Product: metalloproteinase #status experimental <MAT>
 F:428,432/Binding site: zinc, catalytic (His) (active) #status predicted
 F:429/Active site: Glu #status predicted

Query Match 85.4%; Score 35; DB 2; Length 633;
 Best Local Similarity 100.0%; Pred. No. 41;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 NNYRPS 7
 Db 308 NNYRPS 313

RESULT 4
 S42894
 metalloproteinase MEP - Aspergillus fumigatus
 C:Species: Aspergillus fumigatus
 C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Sep-1997
 C:Accession: S42894
 R:Jaton-Ogay, K.; Quadroni, M.; Falchetto, R.; Togni, G.; Paris, S.; Latge, J.; Monod, M.
 submitted to the EMBL Data Library, March 1994
 A:Description: Nucleotide sequence of genomic and cDNA clones encoding an extracellular
 A:Reference number: S42894
 A:Accession: S42894
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-649 <JAT>
 A:Cross-references: EMBL:Z30424; NID:G458475; PID:G458476
 C:Genetics:
 A:Introns: 119/3; 251/2; 588/2

Query Match 85.4%; Score 35; DB 2; Length 649;
 Best Local Similarity 100.0%; Pred. No. 42;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 NNYRPS 7
 Db 308 NNYRPS 313

RESULT 5
 G71607
 probable integral membrane protein PFB0735c - malaria parasite (Plasmodium falciparum)

C:Species: Plasmodium falciparum
 C>Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 09-Jul-2004
 C:Accession: G71607
 R:Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.;
 ; Perte, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.;
 Science 282, 1126-1132, 1998
 A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
 A:Reference number: A71600; MUID:99021743; PMID:9804551
 A:Accession: G71607
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-1182 <GAR>
 A:Cross-references: UNIPROT:O96240; GB:AE001414; GB:AE001362; NID:G3845260; PIDN:AAC7193
 A:Experimental source: clone 3D7
 C:Genetics:
 A:Gene: PFB0735c

Query Match 85.4%; Score 35; DB 2; Length 1182;
 Best Local Similarity 100.0%; Pred. No. 75;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 NNYRPS 7
 Db 624 NNYRPS 629

RESULT 6
 S45889
 probable regulatory protein YBR033w - yeast (Saccharomyces cerevisiae)
 N:Alternate names: hypothetical protein YBR0318
 C:Species: Saccharomyces cerevisiae
 C>Date: 26-Aug-1994 #sequence_revision 09-Sep-1994 #text_change 16-Aug-2004
 C:Accession: S45889; S46562
 R:Grivell, L.A.; de Haan, M.; Maat, M.J.; Smits, P.H.M.
 submitted to the Protein Sequence Database, August 1994
 A:Reference number: S45875
 A:Accession: S45889
 A:Molecule type: DNA
 A:Residues: 1-919 <GRI>
 A:Cross-references: UNIPROT:P38073; EMBL:Z35902; NID:G536247; PIDN:CAA84975.1; PID:G53624
 A:Experimental source: strain S288C
 R:Smits, P.H.M.; de Haan, M.; Maat, C.; Grivell, L.A.
 Yeast 10(Suppl.A), S75-S80, 1994
 A:Title: The complete sequence of a 33 kb fragment on the right arm of chromosome II from
 ly identified genes and a homologue of the SC01 gene.
 A:Reference number: S46551; MUID:94378725; PMID:8091864
 A:Accession: S46562
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-919 <SMI>
 A:Cross-references: EMBL:X76078; NID:G498748; PIDN:CAA53688.1; PID:G498760
 A:Experimental source: strain S288C
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1993
 C:Genetics:
 A:Cross-references: SGD:S0000237
 A:Map position: 2R
 A:Note: YBR033w

C:Superfamily: GAL4 zinc binuclear cluster homology
 C:Keywords: DNA binding; transmembrane protein; zinc finger
 F:51-90/Domain: GAL4 zinc binuclear cluster homology <GAL4>
 F:56-85/Region: zinc finger
 F:408-424/Domain: transmembrane #status predicted <TM1>
 F:482-504/Domain: transmembrane #status predicted <TM2>
 F:547-563/Domain: transmembrane #status predicted <TM3>
 F:703-720/Domain: transmembrane #status predicted <TM4>
 F:790-806/Domain: transmembrane #status predicted <TM5>

Query Match 82.9%; Score 34; DB 2; Length 919;
 Best Local Similarity 85.7%; Pred. No. 93;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
 Db 147 GNNYRPS 153

Db 130 GNNYLPFS 136

RESULT 7

hypothetical protein 311 - diatom (Cylindrotheca fusiformis) plasmid pCF1
C:Species: Cylindrotheca fusiformis
C:Date: 03-May-1994 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C:Accession: S23533
R:Hildebrand, M.; Hasegawa, P.; Ord, R.W.; Thorpe, V.S.; Glass, C.A.; Volcani, B.E.
Plant Mol. Biol. 19, 759-770, 1992
A:Title: Nucleotide sequence of diatom plasmids: identification of open reading frames
A:Reference number: S23531; MUID:92353385; PMID:1322740
A:Accession: S23533
A:Molecule type: DNA
A:Residues: 1-311 <HIL>
A:Cross-references: UNIPROT:Q39500; EMBL:X64302; NID:gl7988; PIDN:CAA45582.1; PID:gl7991
C:Genetics:
C:Genome: plasmid
C:Superfamily: Cylindrotheca fusiformis plasmid pCF1 hypothetical protein 311

Query Match 80.5%; Score 33; DB 2; Length 311;
Best Local Similarity 83.3%; Pred. No. 51;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNYRP 6

Db 108 GNNYQP 113

RESULT 8

sopB protein - Escherichia coli plasmids pO157 and F
N:Alternate names: plasmid partitioning protein; protein B
C:Species: Escherichia coli
C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 09-Jul-2004
C:Accession: T00244; S28096; T42169
R:Makino, K.; Iehi, K.; Yasunaga, T.; Hattori, M.; Yokoyama, K.; Yatsudo, H.C.; Kubota, S.; Shinagawa, H.
DNA Res. 5, 1-9, 1998
A:Title: Complete nucleotide sequences of 93-kb and 3.3-kb plasmids of an enterohemorrhagic E. coli strain O157:H7, substrain RIMD 0509952
A:Reference number: Z14127; MUID:98290540; PMID:9628576
A:Accession: T00244
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-323 <MAK>
A:Cross-references: UNIPROT:P08867; EMBL:AB011549; NID:g4589740; PIDN:BAA31791.1; PID:g4589740; MUID:98290540; PMID:9628576
A:Experimental source: strain EHEC O157:H7, substrain RIMD 0509952
R:Mori, H.; Kondo, A.; Ohshima, A.; Ogura, T.; Hiraga, S.
J. Mol. Biol. 192, 1-15, 1986

A:Title: Structure and function of the F plasmid genes essential for partitioning.

A:Reference number: A92934; MUID:87141188; PMID:3029390

A:Accession: S28096

A:Molecule type: DNA

A:Residues: 1-323 <MOR>

A:Cross-references: EMBL:X04619; NID:g42429; PIDN:CAA28296.1; PID:g42432

R:Burland V.; Shao, Y.; Perna, N.T.; Plunkett, G.; Sofia, H.J.; Blattner, F.R.

Nucleic Acids Res. 26, 4196-4204, 1998

A:Title: The complete DNA sequence and analysis of the large virulence plasmid of Escherichia coli strain O157:H7, substrain RIMD 0509952
A:Reference number: Z22068; MUID:98391744; PMID:9722640

A:Accession: T42169

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-323 <BUR>

A:Cross-references: EMBL:AF074613; PIDN:AAC70137.1

A:Experimental source: strain EDL933; serotype O157:H7

C:Genetics:

A:Gene: sopB

A:Genome: plasmid

C:Superfamily: parB protein

C:Keywords: DNA binding; plasmid partition

Query Match 80.5%; Score 33; DB 2; Length 323;

Best Local Similarity 71.4%; Pred. No. 53;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNYRPS 7

Db 152 GNDYRPT 158

RESULT 9

probable sopB protein - phage N15
N:Alternate names: protein gp27
C:Species: phage N15
C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004
C:Accession: T13113
R:Hendrix, R.W.; Ravin, V.K.; Casjens, S.R.; Ford, M.E.; Ravin, N.V.; Smirnov, I.K.
Submitted to the EMBL Data Library, May 1998
A:Reference number: Z17603
A:Accession: T13113
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-342 <HEN>
A:Cross-references: UNIPROT:O64340; EMBL:AF064539; NID:g3192683; PID:g3192710; PIDN:AAC1111
C:Genetics:
A:Note: gene 27
C:Superfamily: parB protein

Query Match 80.5%; Score 33; DB 2; Length 342;
Best Local Similarity 71.4%; Pred. No. 56;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNYRPS 7

Db 155 GNDYRPT 161

RESULT 10

aminopeptidase A/I XF0138 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: F82843
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequence Database, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: F82843
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-491 <SIM>
A:Cross-references: UNIPROT:Q9PH08; GB:AE003867; NID:g9104919; PIDN:AAF82951
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Albriones, M.R.S.; Bueno, M.R.P.; Camargo, L.E.A.; Camargo, D.M.; Carreir, H. as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
Submitted to GenBank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohme, J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Kriesger, J.E.; Kuramae, E.E.; Laigret, Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E.; F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.; Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak; A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira; M.; Tshukato, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF0138
C:Superfamily: Cytosol aminopeptidase

Query Match 80.5%; Score 33; DB 2; Length 491;
Best Local Similarity 85.7%; Pred. No. 79;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNNYRPS 7

Db 323 GNNYRPS 329

RESULT 11

151556

recombination activating protein - African clawed frog

C:Species: Xenopus laevis (African clawed frog)

C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004

C:Accession: I51556

R:Greenhalgh, P.; Olesen, C.E.; Steiner, L.A.

J. Immunol. 151, 3100-3110, 1993

A:Title: Characterization and expression of recombination activating genes (RAG-1 and RAG-2)

A:Reference number: I51555; MUID:93389137; PMID:8376769

A:Accession: I51556

A>Status: preliminary; translated from GB/EMBL/DDBT

A:Molecule type: DNA

A:Residues: 1-520 <GRE>

A:Cross-references: UNIPROT:Q91830; GB:L19325; NID:G401812; PIDN:AAA49943.1; PID:G401813

C:Genetics:

A:Gene: RAG-2

C:Superfamily: recombination-activating protein RAG-2

Query Match 80.5%; Score 33; DB 2; Length 520;

Best Local Similarity 83.3%; Pred. No. 84;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNYRP 6

Db 225 GNNFRP 230

RESULT 12

F86875

peptide methionine sulfoxide reductase [imported] - Lactococcus lactis subsp. lactis (strain 11337471)

C:Species: Lactococcus lactis subsp. lactis

C>Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004

C:Accession: F86875

R:Botolin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehrlich, S.

Genome Res. 11, 731-753, 2001

A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis subsp. lactis

A:Reference number: A86625; MUID:21235186; PMID:11337471

A:Accession: F86875

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-172 <STO>

A:Cross-references: UNIPROT:Q9CE42; GB:AE005176; PID:G12725052; PIDN:AAK06104.1; GSPDB:G12725052

A:Experimental source: strain IL1403

C:Genetics:

A:Gene: pmsR

C:Superfamily: peptide methionine sulfoxide reductase

Query Match 78.0%; Score 32; DB 2; Length 172;

Best Local Similarity 83.3%; Pred. No. 45;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNYRP 6

Db 92 GNNYRP 97

RESULT 13

A35009

C 3.4.23.42 thermopsin () precursor - Sulfolobus acidocaldarius

C:Species: Sulfolobus acidocaldarius

C>Date: 20-Jul-1990 #sequence_revision 20-Jul-1990 #text_change 09-Jul-2004

C:Accession: A35009

R:Lin, X.; Tang, J.

J. Biol. Chem. 265, 1490-1495, 1990

A:Title: Purification, characterization, and gene cloning of thermopsin, a thermostable

A:Reference number: A35009; MUID:90110209; PMID:2104843

A:Accession: A35009

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-340 <LIN>

A:Cross-references: UNIPROT:P17118; GB:J05184; NID:G152940; PID:G152941

C:Superfamily: Sulfolobus acidocaldarius thermopsin

C:Keywords: hydrolase

Query Match 78.0%; Score 32; DB 2; Length 340;

Best Local Similarity 71.4%; Pred. No. 88;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNNYRPS 7

Db 246 GNNYTPN 252

RESULT 14

B46216

transcription factor Skn-1, splice form i - rat

N:Alternate names: Skn-1i

C:Species: Rattus norvegicus (Norway rat)

C>Date: 19-Nov-1993 #sequence_revision 18-Nov-1994 #text_change 05-Feb-1999

C:Accession: B46216

R:Andersen, B.; Schonemann, M.D.; Flynn, S.E.; Pearce II, R.V.; Singh, H.; Rosenfeld, M.C.

Science 260, 78-82, 1993

A:Title: Skn-1a and Skn-1i: two functionally distinct Oct-2-related factors expressed in

A:Reference number: A46216; MUID:93219836; PMID:7682011

A:Accession: B46216

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-348 <AND>

A:Cross-references: GB:L23863; NID:G393221

A>Note: sequence extracted from NCBI backbone (NCBIP:128574)

C:Superfamily: transcription factor Oct-2; homeobox homology; POU domain homology

C:Keywords: alternative splicing; DNA binding; homeobox; nucleus; transcription regulation

F:101-168/Domain: POU domain homology <POU>

F:193-249/Domain: homeobox homology <HOX>

Query Match 78.0%; Score 32; DB 1; Length 348;

Best Local Similarity 85.7%; Pred. No. 90;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNNYRPS 7

Db 297 GNNSRPS 303

RESULT 15

T49791

hypothetical protein B9J10.280 [imported] - Neurospora crassa

C:Species: Neurospora crassa

C>Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004

C:Accession: T49791

R:Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura, J.

submitted to the Protein Sequence Database, May 2000

A:Reference number: Z25022

A:Accession: T49791

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-349 <SCH>

A:Cross-references: UNIPROT:Q9P555; EMBL:AL356324; GSPDB:GN00116; NCSP:B9J10.280

A:Experimental source: BAC clone B9J10; strain OR74A

C:Genetics:

A:Gene: NCSP:B9J10.280

A:Map position: 6

Query Match 78.0%; Score 32; DB 2; Length 349;

Best Local Similarity 83.3%; Pred. No. 90;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 NNVRPS 7

Db |||||:
 145 NNYRPA 150

Search completed: April 28, 2005, 18:25:51
Job time : 11.625 secs

This Page Blank (uspiu)

GenCore version 5.1.6

Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 01:42:46 ; Search time 5.81132 Seconds
(without alignments)
62.721 Million cell updates/sec

Title: US-10-088-639A-2_COPY_49_55

Perfect score: 41

Sequence: 1 GNNYRPS 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	37	90.2	951	1	XRN2_MOUSE
2	36	87.8	387	1	SSUD_BRAJA
3	35	85.4	222	1	T2M1_METJA
4	35	85.4	634	1	ELM1_ASPPU
5	35	85.4	634	1	ELM2_ASPPU
6	34	82.9	919	1	YB03_YEAST
7	33	80.5	323	1	SOPB_ECOLI
8	33	80.5	349	1	HS2S_DROME
9	33	80.5	490	1	AMPA_XANAC
10	33	80.5	491	1	AMPA_XYLFA
11	33	80.5	491	1	AMPA_XYLFT
12	33	80.5	493	1	AMPA_XANCP
13	33	80.5	520	1	RAG2_XENLA
14	33	80.5	581	1	HPLC_MOUSE
15	33	80.5	847	1	CADQ_HUMAN
16	32	78.0	169	1	MSRA_STRP3
17	32	78.0	169	1	MSRA_STRP8
18	32	78.0	169	1	MSRA_STRPY
19	32	78.0	172	1	MSA2_LACLA
20	32	78.0	340	1	THPS_SULAC
21	32	78.0	417	1	ARPA_HUMAN
22	32	78.0	430	1	PO23_RAT
23	32	78.0	431	1	PO23_MOUSE
24	32	78.0	436	1	PO23_HUMAN
25	32	78.0	707	1	Z224_HUMAN
26	32	78.0	720	1	MR11_ARATH
27	32	78.0	1097	1	IMB3_HUMAN
28	32	78.0	1914	1	STCK_EMENI
29	31	75.6	92	1	RM34_HUMAN
30	31	75.6	92	1	RM34_MOUSE
31	31	75.6	105	1	RM34_YEAST
32	31	75.6	108	1	RM34_SCHPO
33	31	75.6	236	1	KDGM_ERWCH
34	31	75.6	31	1	CAH2_TRIHK
35	31	75.6	31	1	DPH5_SCHPO
36	31	75.6	31	1	CAH5_MOUSE
37	31	75.6	31	1	CAH5_HUMAN
38	31	75.6	31	1	KDGM_ECOLI
39	31	75.6	31	1	KDGM_ERWCH
40	31	75.6	31	1	EFTU_RECAM
41	31	75.6	31	1	EFTU_CYTLY
42	31	75.6	31	1	EFTU_LISIN
43	31	75.6	31	1	EFTU_LISMO
44	31	75.6	31	1	EFTU_TAXOC
45	31	75.6	31	1	EFTU_BURCE
46	31	75.6	31	1	EFTU_MYCLE
47	31	75.6	31	1	EFTU_MYCTU
48	31	75.6	31	1	EFTU_RALSO
49	31	75.6	31	1	EFTU_THICU
50	31	75.6	31	1	EFT1_STRCO
51	31	75.6	31	1	EFT1_STRCU
52	31	75.6	31	1	EFT1_STRRA
53	31	75.6	31	1	EFTU_PLARO
54	31	75.6	31	1	EFTU_STRAU
55	31	75.6	31	1	EFTU_STRCJ
56	31	75.6	31	1	EFTU_STRMU
57	31	75.6	31	1	EFTU_STRP3
58	31	75.6	31	1	EFTU_STRP8
59	31	75.6	31	1	EFTU_STRPN
60	31	75.6	31	1	EFTU_STRPY
61	31	75.6	31	1	EFTU_CANJE
62	31	75.6	31	1	CYCL_DROME
63	31	75.6	31	1	OPGG_ECOL6
64	31	75.6	31	1	OPGG_ECOLI
65	31	75.6	31	1	OPGG_SHIFL
66	31	75.6	31	1	ABEI_HUMAN
67	31	75.6	31	1	CLPE_STRPN
68	31	75.6	31	1	POL1_BAYMJ
69	31	75.6	31	1	POL1_BAYMG
70	31	75.6	31	1	HV60_MOUSE
71	31	75.6	31	1	DIIP_MOUSE
72	31	75.6	31	1	TTPA_HUMAN
73	31	75.6	31	1	YDCL_SCHPO
74	31	75.6	31	1	HEMZ_RHOCA
75	30	73.2	226	1	FMA5_PORGI
76	30	73.2	278	1	INU2_ASPFI
77	30	73.2	278	1	SMF2_YEAST
78	30	73.2	297	1	RPOC_VARPO
79	30	73.2	351	1	CD96_DROME
80	30	73.2	367	1	MCR_HUMAN
81	30	73.2	396	1	VGL2_CVMJH
82	30	73.2	516	1	VGL2_CVM4
83	30	73.2	549	1	VGL2_CVMJC
84	30	73.2	684	1	IP3S_RAT
85	30	73.2	773	1	IP3S_MOUSE
86	30	73.2	984	1	MSA1_LACLA
87	30	73.2	1235	1	MSUE_PSEAE
88	30	73.2	1376	1	Y476_AERPE
89	30	73.2	1376	1	LB31_ARATH
90	30	73.2	2701	1	MOVV_TAV
91	29	70.7	120	1	MOVV_PSVJ
92	29	70.7	183	1	CAH8_HUMAN
93	29	70.7	186	1	CAH8_MOUSE
94	29	70.7	197	1	UL14_HSVB
95	29	70.7	215	1	YDPA_BACSU
96	29	70.7	277	1	YDPA_BACSU
97	29	70.7	288	1	E133_SOLUTU
98	29	70.7	289	1	E131_SOLUTU
99	29	70.7	290	1	KDGR_BACSU
100	29	70.7	317	1	E13B_PHAVU
101	29	70.7	324	1	SXL_GERCA
102	29	70.7	328	1	CAH8_MOUSE
103	29	70.7	337	1	CAH8_MOUSE
104	29	70.7	339	1	CAH8_MOUSE
105	29	70.7	348	1	CAH8_MOUSE
106	29	70.7	348	1	CAH8_MOUSE

Q8uwa5 tribolodon
O74898 schizosacch
P23589 mus musculu
P15218 homo sapien
P37647 escherichia
P45416 erwinia chr
O21245 reclinomona
P42474 cytophaga l
O92716 listeria in
O8Y422 listeria mo
P42480 taxobacter
P33167 burkholderi
P30768 mycobacteri
P31501 mycobacteri
O8XG20 ralstonia s
P42481 thiobacillu
P40174 streptomyc
Q53871 streptomyc
P29542 streptomyc
P72231 planobisp
O33594 streptomyc
P95724 streptomyc
P72483 streptomyc
P33170 streptomyc
O8K872 streptomyc
O97Pv3 streptomyc
P82559 streptomyc
O69303 campylobact
O61734 drosophila
P02992 saccharomyc
Q8Cw60 escherichia
P31336 escherichia
O83ru3 shigella fl
Q96bi0 homo sapien
P35594 streptomyc
P07168 agrobacteri
Q4584 human immun
O01206 barley yell
O04574 barley yell
P18531 mus musculu
Q9Dca7 mus musculu
P49638 homo sapien
P41034 rattus norv
Q10422 schizosacch
O59735 rhodobacter
Q9S0w8 porphyromon
P33232 escherichia
O94220 aspergillus
P38778 saccharomyc
P06273 marchantia
Q9Vbw3 drosophila
P08235 homo sapien
P11225 murine coro
P22432 murine coro
Q02385 murine coro
P29955 rattus norv
O24418 gossypium h
Q9Cfc8 lactococcus
O31038 pseudomonas
Q9Yev5 aeropyrum p
O81322 arabidopsis
P23226 tomato aspe
P22117 peanut stun
P35219 homo sapien
P28651 mus musculu
P28949 equine herp
P50736 bacillus su
P52402 solanum tub
P52400 solanum tub
P50844 bacillus su
P23535 phaseolus v
O61374 ceratitidis c

```

107 29 70.7 360 1 E13B LYCES
108 29 70.7 363 1 E132 SOLTU
109 29 70.7 365 1 E13E NICPL
110 29 70.7 370 1 E13E TOBAC
111 29 70.7 371 1 E13F TOBAC
112 29 70.7 371 1 E13F TOBAC
113 29 70.7 372 1 3B4 RAT
114 29 70.7 373 1 YF4 YFAST
115 29 70.7 374 1 E13B HEVR
116 29 70.7 380 1 BCAL HELPJ
117 29 70.7 381 1 RODA HELPJ
118 29 70.7 385 1 AMPH ECOLI
119 29 70.7 417 1 ARPA MOUSE
120 29 70.7 436 1 RHO AQAE
121 29 70.7 491 1 AMPA HAEN
122 29 70.7 493 1 AMPA AQAE
123 29 70.7 493 1 HMB DROME
124 29 70.7 494 1 AMPA PASMU
125 29 70.7 500 1 AMAL SHEON
126 29 70.7 502 1 AMA2 SHEON
127 29 70.7 533 1 LEUI MITCAE
128 29 70.7 538 1 YQ4 CANAL
129 29 70.7 557 1 YKPE YFAST
130 29 70.7 628 1 TOP1 ARCFU
131 29 70.7 653 1
132 29 70.7 679 1 TKTI KUULA
133 29 70.7 769 1 SWI6 KUULA
134 29 70.7 937 1 PAC4 RAT
135 29 70.7 958 1 YAB7 SCHPO
136 29 70.7 968 1 IRS1 DROME
137 29 70.7 969 1 PAC4 HUMAN
138 29 70.7 1005 1 MANA DICDI
139 29 70.7 1250 1 BXE CLOBO
140 29 70.7 1250 1 MYOF HUMAN
141 29 70.7 2061 1
142 29 70.7 5127 1 RY44 DROME
143 28 68.3 208 1 Y196 THEAC
144 28 68.3 222 1 SDHE PEPAS
145 28 68.3 247 1 RLUB BUCBP
146 28 68.3 258 1 ARAS ARATH
147 28 68.3 266 1 RFA2 RAT
148 28 68.3 270 1 RFA2 MOUSE
149 28 68.3 285 1 ROAM MOUSE
150 28 68.3 287 1 ARYL CHICK

```

ALIGNMENTS

```

RESULT 1
XR2N2_MOUSE STANDARD; PRT; 951 AA.
ID XR2N2_MOUSE Q99K57;
AC Q9DBR1; Q61489; Q99K57;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE 5'-3', exoribonuclease 2 (EC 3.1.11.-) (Dhml protein).
GN XR2N2 OR DHM1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Testis;
RX MEDLINE=95192042; PubMed=7885830;
RA Shobuika T., Sugano S., Yamashita T., Ikeda H.;
RT "Characterization of cDNA encoding mouse homolog of fission yeast
RT dhpl+ gene: structural and functional conservation.";
RL Nucleic Acids Res. 23:357-361(1995).
RN [2]
RN SEQUENCE FROM N.A. (ISOFORM 1).
RP STRAIN=C57BL/6J; TISSUE=Embryo, Eye, Forelimb, and Lung;
RC

```

```

RX MEDLINE=22354683; PubMed=12466851;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojibori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.W.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chochia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gusincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Nunata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Sempile C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita K.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavalan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RA "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [3]
RN SEQUENCE FROM N.A. (ISOFORM 1)
RP STRAIN=C57BL/6, and Czech II; TISSUE=Brain, and Mammary gland;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner K.H., Shenmen C.M., Schuler G.D.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ussidi T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Possesses 5'->3' exoribonuclease activity and may be
CC involved in homologous recombination and RNA metabolism, such as
CC RNA synthesis and RNA trafficking.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q9DBR1-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q9DBR1-2; Sequence=VSP_007235;
CC Note=No experimental confirmation available. May result from the
CC retention of an intron in the cDNA.
CC -!- TISSUE SPECIFICITY: Expressed in the spleen, testis, heart, brain,
CC lung, liver, skeletal muscle, and kidney.
CC -!- SIMILARITY: Belongs to the 5'-3' exonuclease family.
CC -!- SIMILARITY: Contains 1 CCHC-type zinc finger.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration -
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```

CC  EMBL; D38517; BAA07524.1; -.
CC  EMBL; AK004800; BAB23573.1; -.
CC  EMBL; AK031247; BAC27318.1; -.
CC  EMBL; AK053643; BAC35458.1; -.
CC  EMBL; BC004028; AAH04028.1; ALT_INIT.
CC  EMBL; BC054743; AAH54743.1; -.
CC  MGD; MGI:894687; Xrn2.
CC  InterPro; IPR004859; Put 53exo.
CC  InterPro; IPR001878; Znf_CCHC.
CC  Pfam; PF03159; XRN_N; 1.
CC  SMART; SM00343; Znf_C2HC; 1.
CC  PROSITE; PS0158; ZF_CCHC; FALSE NEG.
KW  mRNA processing; Hydrolase; Exonuclease; Nuclear protein;
KW  RNA-binding; Zinc-finger; Alternative splicing.
FT  ZN FING 262 278
FT  VARSPLOC 931 951
FT  C (in isoform 2).
FT  /FTID=VSP_007235.
FT  SKE -> IKG (IN REF. 1).
FT  H -> Q (IN REF. 1).
FT  I -> N (IN REF. 1).
FT  V -> E (IN REF. 1).
FT  K -> R (IN REF. 1).
FT  Y -> L (IN REF. 1).
FT  P -> H (IN REF. 2; BAC27318).
FT  T -> K (IN REF. 2; BAC27318).
FT  P -> L (IN REF. 1).
FT  Q -> K (IN REF. 1).
FT  P -> H (IN REF. 2; BAC27318).
FT  Q -> QV (IN REF. 2; BAC27318).
SQ  SEQUENCE 951 AA; 108687 MW; CF57479291DD18B9 CRC64;

```

Query Match 90.2%; Score 37; DB 1; Length 951;
 Best Local Similarity 100.0%; Pred. No. 7.9;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GNNYRP 6
DB 843 GNNYRP 848

```

```

RESULT 2
SSUD_BRAJA STANDARD; PRT; 387 AA.
AC Q89ER2;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Alkanesulfonate monooxygenase (EC 1.14.14.5) (FMNH2-dependent
DE aliphatic sulfonate monooxygenase).
GN SSUD OR BL27010.
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]_SEQUENCE FROM N.A.
RP STRAIN=USDA 110;
RC MEDLINE=22484998; PubMed=12597275;
RX Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Sasamoto S., Watanabe A., Idesawa K., Iriuchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110."
RL DNA Res. 9:189-197(2002).
CC -!- FUNCTION: Catalyzes the desulfonation of aliphatic sulfonates (BY

```

```

CC similarity).
CC -!- CATALYTIC ACTIVITY: An alkanesulfonate (R-CH(2)-SO(3)H) + FMNH(2) +
CC O(2) = an aldehyde (R-CHO) + FMN + sulfite + H(2)O.
CC -!- SIMILARITY: Belongs to the ssud family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).


Query Match 87.8%; Score 36; DB 1; Length 387;  

  Best Local Similarity 85.7%; Pred. No. 5.1;  

  Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;



```

QY 1 GNNYRPS 7
DB 375 GNDYRPS 381

```



RESULT 3  

  T2M1_METJA STANDARD; PRT; 222 AA.  

  ID T2M1_METJA  

  AC Q58391;  

  DT 01-NOV-1997 (Rel. 35, Created)  

  DT 01-NOV-1997 (Rel. 35, Last sequence update)  

  DT 10-OCT-2003 (Rel. 42, Last annotation update)  

  DE Type II restriction enzyme Mjai (EC 3.1.21.4) (Endonuclease Mjai)  

  DE (R.Mjai).  

  GN MJAIR OR MJ0984.  

  OS Methanococcus jannaschii.  

  OC Archaea; Euryarchaeota; Methanococci; Methanococcales;  

  OC Methanocaldococcaceae; Methanocaldococcus.  

  OX NCBI_TaxID=2190;  

  RN [1]  

  RP SEQUENCE FROM N.A.  

  RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;  

  RX MEDLINE=96337999; PubMed=8689087;  

  RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,  

  RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,  

  RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,  

  RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,  

  RA Scott J.L., Geoghegan N.S.M., Weidman J.P., Fuhrmann J.L., Nguyen D.,  

  RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,  

  RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,  

  RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;  

  RT "Complete genome sequence of the methanogenic archaeon, Methanococcus  

  RT jannaschii.";  

  RL Science 273:1058-1073(1996).  

  RN [2]  

  RC CHARACTERIZATION.  

  RP Noren C.J., Roberts R.J., Patti J., Byrd D.R., Morgan R.D.;  

  RT "Method for screening restriction endonucleases.";  

  RT Patent number WO9911821, 11-MAR-1999.  

  CC -!- FUNCTION: RECOGNIZES THE DOUBLE-STRANDED SEQUENCE CTAG  

  CC AND CLEAVES AFTER C-1 (BY SIMILARITY).  

  CC -!- CATALYTIC ACTIVITY: Endonucleolytic cleavage of DNA to give  

  CC 'specific double-stranded fragments with terminal 5'-phosphates.  

  CC -!- SIMILARITY: STRONG, TO M.THERMOFORMICICUM MTHZIR.  

  CC -----
  CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
  CC the European Bioinformatics Institute. There are no restrictions on its
  CC use by non-profit institutions as long as its content is in no way


```

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; U67541; AAB98987.1; -.
 DR PIR; H64422; H64422.
 DR REBASE; 1221; Mjai.
 DR TIGR; M0984; -.
 KW Restriction system; Hydrolase; Nuclease; Endonuclease;
 KW Complete proteome.
 SQ SEQUENCE 222 AA; 25962 MW; 9A254A208AB8AC02 CRC64;

Query Match 85.4%; Score 35; DB 1; Length 222;

Best Local Similarity 85.7%; Pred. No. 4.6;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNNYRPS 7

Db 147 GGNYPRES 153

RESULT 4

ELM1 ASPFU
 ID ELM1 ASPFU STANDARD; PRT; 634 AA.
 AC P4607A;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Extracellular elastinolytic metalloproteinase precursor (BC 3.4.24.-).
 GN MEP.

OS Aspergillus fumigatus (Sartorya fumigata).

OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;

OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.

OX NCBI_TaxID=5085;

RN [1] SEQUENCE FROM N.A., AND SEQUENCE OF 530-542.

RC STRAIN=Isolate 13;

RX MEDLINE=95012603; PubMed=7927676;

RA Sirakova T.D., Markaryan A., Kolattukudy P.E.;

RT "Molecular cloning and sequencing of the cDNA and gene for a novel

RT elastinolytic metalloproteinase from Aspergillus fumigatus and its

RT expression in Escherichia coli.";

RL Infect. Immun. 62:4208-4218(1994).

RN [2]

RP SEQUENCE OF 246-258.

RX MEDLINE=94245315; PubMed=8188335;

RA Markaryan A., Morozova I., Yu H., Kolattukudy P.E.;

RT "Purification and characterization of an elastinolytic

RT metalloproteinase from Aspergillus fumigatus and immunoelectron

RT microscopic evidence of secretion of this enzyme by the fungus

RT invading the murine lung.";

RL Infect. Immun. 62:2149-2157(1994).

CC 1- FUNCTION: Catalyzes the hydrolysis of elastin.

CC 1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).

CC 1- SIMILARITY: Belongs to peptidase family M36.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial/
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; L29566; AAB07708.1; -.

DR MEROPS; M36.001; -.

DR InterPro; IPR006025; Pept_M_Zn_BS.

DR InterPro; IPR001842; Peptidase M36.

DR Pfam; PF02128; Peptidase M36; 1.

DR PRINTS; PR00999; FUNGALYSIN.

DR PROSITE; PS00142; ZINC_PROTEASE; 1.

KW Hydrolase; Metalloproteinase; Zinc; Signal; Zymogen.

FT SIGNAL 1 18 POTENTIAL.

FT PROPEP 19 245
 FT CHAIN 246 634
 FT -----
 FT METAL 429 429
 FT ACT SITE 430 430
 FT METAL 433 433
 FT METAL 433 433
 SQ SEQUENCE 634 AA; 68726 MW; B94E274BF767F911 CRC64;

Query Match 85.4%; Score 35; DB 1; Length 634;

Best Local Similarity 100.0%; Pred. No. 14;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 NNYRPS 7

Db 309 NNYRPS 314

RESULT 5

ELM2 ASPFU

ID ELM2 ASPFU STANDARD; PRT; 634 AA.

AC P46075;

DT 01-NOV-1995 (Rel. 32, Created)

DT 15-DEC-1998 (Rel. 37, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Extracellular elastinolytic metalloproteinase precursor (BC 3.4.24.-).

GN MEP.

OS Aspergillus fumigatus (Sartorya fumigata).

OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;

OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.

OX NCBI_TaxID=5085;

RN [1] SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

RC STRAIN=DELTA18;

RX MEDLINE=95231298; PubMed=7715453;

RA Jaton-Ogay K., Paris S., Huerre M., Quadroni M., Falchetto R.,

RA Togni G., Latge J.-P., Monod M.;

RT "Cloning and disruption of the gene encoding an extracellular

RT metalloproteinase of Aspergillus fumigatus.";

RL Mol. Microbiol. 14:917-928(1994).

RN [2]

RP REVISIONS.

RC STRAIN=DELTA18;

RA Sanglard D.;

RL Submitted (OCT-1998) to the EMBL/Genbank/DBJ databases.

CC 1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).

CC 1- SIMILARITY: Belongs to peptidase family M36.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial/
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; Z30424; CAA83015.1; -.

DR MEROPS; M36.001; -.

DR InterPro; IPR006025; Pept_M_Zn_BS.

DR InterPro; IPR001842; Peptidase M36.

DR Pfam; PF02128; Peptidase M36; 1.

DR PRINTS; PR00999; FUNGALYSIN.

DR PROSITE; PS00142; ZINC_PROTEASE; 1.

KW Hydrolase; Metalloproteinase; Zinc; Signal; Zymogen.

FT SIGNAL 1 18 POTENTIAL.

FT PROPEP 19 245

FT CHAIN 246 634

FT METAL 429 429

FT ACT SITE 430 430

FT METAL 433 433

FT METAL 433 433

SQ SEQUENCE 634 AA; 68718 MW; 7BA5A9B70DE184BE CRC64;

Query Match 85.4%; Score 35; DB 1; Length 634;

```

Best Local Similarity 100.0%; Pred. No. 14;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 NNYRPS 7
Db      309 NNYRPS 314
|||||

RESULT 6
YB03 YEAST
ID YB03 YEAST STANDARD; PRT; 919 AA.
AC P38073;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative 103.4 kDa transcriptional regulatory protein in RPLA4A-HWT1
DE intergenic region
GN YB033W OR YB03318.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c;
RX MEDLINE=94378725; PubMed=8091864;
RA Smite P.H.M., de Haan M., Maat C., Grivell L.A.;
RT "The complete sequence of a 33 kb fragment on the right arm of
RT chromosome II from Saccharomyces cerevisiae reveals 16 open reading
RT frames, including ten open reading frames, five previously
RT identified genes and a homologue of the SCO1 gene.";
RL Yeast 10:575-580(1994).
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- SIMILARITY: Contains 1 Zn(2)-Cys(6) fungal-type binuclear cluster
CC domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X76078; CAA53688.1; -
DR EMBL; Z35902; CAA84975.1; -
DR FIR; S45889; S45889.
DR Germline; 138576; -.
DR SGD; S0000237; YBR033W.
DR InterPro; IPR001138; Fungi_Trn.
DR Pfam; PF00172; Zn.clus; 1.
DR SMART; SM00066; GFL4; 1.
DR PROSITE; PS00463; ZN2_Cy6_FUNGAL_1; 1.
DR PROSITE; PS00448; ZN2_Cy6_FUNGAL_2; 1.
KW Hypothetical protein; Transcription regulation; DNA-binding;
KW Nuclear protein; Zinc; Metal-binding.
FT DNA_BIND 56 85 ZN(2)-CYS(6), FUNGAL-TYPE.
SQ SEQUENCE 919 AA; 103396 MW; 18DD2E37A042E246 CRC64;

Query Match 82.9%; Score 34; DB 1; Length 919;
Best Local Similarity 85.7%; Pred. No. 32;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GNNYRPS 7
Db      130 GNNYLPFS 136
|||||

RESULT 7
IDB_SCOLI
ID SOPB_ECOLI STANDARD; PRT; 323 AA.
AC P08867;
DT 01-NOV-1988 (Rel. 09, Created)

```

```

DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE SopB protein (Plasmid partition protein B).
GN SOPB OR B.
OS Escherichia coli, and
OS Escherichia coli O157:H7.
OG Plasmid F, and Plasmid pO157.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562, 83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12; PLASMID=F;
RX MEDLINE=87141188; PubMed=3029390;
RA Mori H., Kondo A., Ohshima A., Ogura T., Hiraga S.;
RT "Structure and function of the F plasmid genes essential for
RT partitioning.";
RL J. Mol. Biol. 192:1-15(1986).
RN [2]
RP SEQUENCE FROM N.A.
RC PLASMID=F;
RA Eichenlaub R.;
RT "F plasmid DNA complete mini-F region (F coordinates 40.301F to
RT 49.869F).";
RL Submitted (AUG-1986) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / CR63; PLASMID=F;
RA Shmizu H., Saitho Y., Suda Y., Uehara K., Sampei G., Mizobuchi K.;
RT "Complete nucleotide sequence of the F plasmid: its implications for
RT organization and diversification of plasmid genomes.";
RL Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / R1MD 0509952; PLASMID=pO157;
RX MEDLINE=98290540; PubMed=9628576;
RA Makino K., Ishii K., Yasunaga T., Hattori M., Yokoyama K.,
RA Yutsudo H.C., Kubota Y., Yamauchi Y., Iida T., Yamamoto K., Honda T.,
RA Han C.G., Ohtsubo E., Kasamatsu M., Hayaashi T., Kuhara S.,
RA Shinagawa H.;
RT "Complete nucleotide sequences of 93-kb and 3.3-kb plasmids of an
RT enterohemorrhagic Escherichia coli O157:H7 derived from Sakai
RT outbreak.";
RL DNA Res. 5:1-9(1998).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927; PLASMID=pO157;
RX MEDLINE=98391744; PubMed=9722640;
RA Burland V., Shao Y., Perna N.T., Plunkett G., Sofia H.J.,
RA Blattner F.R.;
RT "The complete DNA sequence and analysis of the large virulence plasmid
RT of Escherichia coli O157:H7.";
RL Nucleic Acids Res. 26:4196-4204(1998).
CC -!- FUNCTION: CONTROL OF PLASMID PARTITIONING; REQUIRED TO RECOGNIZE
CC THE CYS-ACTING. BINDS SPECIFICALLY WITH THE DNA SEGMENT CONTAINING
CC THE SOPC REGION. SOPB IS TRANS-ACTING.
CC -!- MISCELLANEOUS: OVERPRODUCTION OF SOPB PROTEIN CAUSES INCC
CC INCOMPATIBILITY.
CC -!- SIMILARITY: Belongs to the parB family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X04619; CAA28296.1; -
DR EMBL; M12987; AAA24903.1; -
DR EMBL; AP001918; BAA97917.1; -
DR EMBL; AB011549; BAA31791.1; -
DR EMBL; AF074613; AAC70137.1; -

```

```

DR PIR: T00244; T00244.
DR EcoGene; EG40060; sopB.
DR InterPro; IPR004437; ParB part.
DR InterPro; IPR003115; ParBc.
DR Pfam; PF02195; ParB; 1.
DR SMART; SM00470; ParB; 1.
DR TIGRFAMs; TIGR00180; parB part; 1.
KW Plasmid; Plasmid partition; DNA-binding; Complete proteome.
FT DNA_BIND 180 199 H-T-H MOTIF (BY SIMILARITY).
SQ SEQUENCE 323 AA; 35372 MW; 3CC19C873CED3C1C CRC64;

Query Match 80.5%; Score 33; DB 1; Length 323;
Best Local Similarity 71.4%; Pred. No. 18;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
Db 152 GNDYRPT 158
:::|:::

RESULT 8
HS2S DROME STANDARD; PRT; 349 AA.
AC P25722; Q9VIM4;
DT 01-MAY-1992 (Rel. 22, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE Heparin sulfatase O-sulfotransferase (EC 2.8.2.-).
GN HS2ST OR CG10234.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC STRAIN=Oregon-R;
RX MEDLINE=92038937; PubMed=1936954;
RA Powers P.A., Ganetzky B.;
RT "On the components of segregation distortion in Drosophila
RT melanogaster. V. Molecular analysis of the sd locus.";
RL Genetics 129:133-144 (1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abrell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotter P.,
RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA De Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

```

```

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-P., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195 (2000).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM S1).
RC STRAIN=Berkely; TISSUE=Head;
RX MEDLINE=22426066; PubMed=12537569;
RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,
RA George R.A., Guarin H., Kronmiller B., Pacle J.M., Park S., Wan K.H.,
RA Rubin G.M., Celniker S.E.;
RT "A Drosophila full-length cDNA resource.";
RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8 (2002).
CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Golgi membrane (By
CC similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Comment=Additional isoforms seem to exist;
CC Name=S1;
CC IsoId=P25722-1; Sequence=Displayed;
CC Name=S2;
CC IsoId=P25722-2; Sequence=VSP_004382;
CC Name=S4;
CC IsoId=P25722-3; Sequence=VSP_004383;
CC -!- CAUTION: Was originally (Ref.1) thought to be SD.
CC -!- CAUTION: Ref.1 sequence differs from that shown due to a
CC frameshift in position 296.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X60218; CA942779.1; ALT_FRAME.
DR EMBL; AE003663; AAF53800.1; -.
DR EMBL; AY058422; AAL13651.1; -.
DR PIR; S18765; S18765.
DR FlyBase; FBgn0024230; Hs2st.
DR InterPro; IPR007734; HS2ST.
DR Pfam; PF05040; HS2ST; 1.
KW Transferrase; Transmembrane; Glycoprotein; Golgi stack; Signal-anchor;
FT DOMAIN 1 17
FT TRANSMEM 18 38
FT CYTOPLASMIC (POTENTIAL).
FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT (POTENTIAL)
FT LUMENAL (POTENTIAL).
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT MFRLKMWLLRPHLTHLILALC -> MKRGAECSEMQAF
FT PESDDGFRQPGIITIDEAFEAI (in isoform S2).
FT /FTID=VSP_004382.
FT Missing (in isoform S4).
FT /FTID=VSP_004383.
FT C -> S (IN REF. 1).
FT DN -> EH (IN REF. 1).
FT CONFLICT 28 28
FT CONFLICT 180 181
FT SEQUENCE 349 AA; 41273 MW; E4655D92D7615C41 CRC64;
SQ
Query Match 80.5%; Score 33; DB 1; Length 349;
Best Local Similarity 71.4%; Pred. No. 19;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 GNNYRPS 7

```

```
Db 179 GDNYPN 185
|:|||||
RESULT 9
ID AMPA_XANAC STANDARD; PRT; 490 AA.
AC Q8PGR0;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable cytosol aminopeptidase (EC 3.4.11.1) (Leucine aminopeptidase)
DE (LAP) (Leucyl aminopeptidase)
GN PEPA OR XAC3556.
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=92829;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=22022145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Cannavan F., Cardozo J., Chamberg F., Clapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira L.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola H.A.F., Takita M.A., Tamura R.B., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."
RL Nature 417:459-463(2002).
CC -!- FUNCTION: Presumably involved in the processing and regular
CC turnover of intracellular proteins. Catalyzes the removal of
CC unsubstituted N-terminal amino acids from various peptides (By
CC similarity).
CC -!- CATALYTIC ACTIVITY: Release of an N-terminal amino acid, Xaa-|-
CC including Pro although not Arg or Lys, and Xbb may be other amino acids
CC -!- COFACTOR: Binds 2 manganese ions per subunit (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to peptidase family M17.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF012005; AAM38399.1; -.
CC HAMAP; MF 00181; -.
CC InterPro; IPR000819; Peptidase_M17_C.
CC InterPro; IPR008283; Peptidase_M17_N.
CC Pfam; PF00883; Peptidase_M17_1.
CC Pfam; PF02789; Peptidase_M17_N; 1.
CC PRINTS; PR00481; LAMNOPPTDASE.
CC PROSITE; PS00631; CYTOSOL AP; 1.
CC Hydrolase; Aminopeptidase; Manganese; Complete proteome.
FT ACT_SITE 274 274 POTENTIAL.
FT ACT_SITE 348 348 POTENTIAL.
FT METAL 262 267 MANGANESE 2 (BY SIMILARITY).
FT METAL 267 267 MANGANESE 1 AND 2 (BY SIMILARITY).
FT METAL 285 285 MANGANESE 2 (BY SIMILARITY).
FT METAL 285 285 MANGANESE 2 (BY SIMILARITY).

FT METAL 344 344 MANGANESE 1 (BY SIMILARITY).
FT METAL 346 346 MANGANESE 1 AND 2 (BY SIMILARITY).
SQ SEQUENCE 490 AA; 51181 MW; E3D2A310638965E9 CRC64;

Query Match 80.5%; Score 33; DB 1; Length 490;
Best Local Similarity 85.7%; Pred. No. 27;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
DB 321 GNNYRPS 327

RESULT 10
ID AMPA_XYLFA STANDARD; PRT; 491 AA.
AC Q9PH08;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable cytosol aminopeptidase (EC 3.4.11.1) (Leucine aminopeptidase)
DE (LAP) (Leucyl aminopeptidase).
GN PEPA OR XF0138.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xylella.
OX NCBI_TaxID=2371;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=9a5C;
RX MEDLINE=20365717; PubMed=10910347;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvares R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carver H.,
RA Coutinho L.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hohnselt J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuramae E.B., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A. Jr., Nobrega P.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A., Terenzi M.F., Truffi D., Tsai S.M., Tshahko M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa."
RL Nature 406:151-159(2000).
CC -!- FUNCTION: Presumably involved in the processing and regular
CC turnover of intracellular proteins. Catalyzes the removal of
CC unsubstituted N-terminal amino acids from various peptides (By
CC similarity).
CC -!- CATALYTIC ACTIVITY: Release of an N-terminal amino acid, Xaa-|-
CC Xbb-, in which Xaa is preferably Leu, but may be other amino acids
CC including Pro although not Arg or Lys, and Xbb may be Pro.
CC -!- COFACTOR: Binds 2 manganese ions per subunit (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to peptidase family M17.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
```


RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
 RA Setubal J.C., Kitajima J.P.;
 RT "Comparison of the genomes of two Xanthomonas pathogens with differing
 RT host specificities.";
 RL Nature 417:459-463(2002).
 CC -!- FUNCTION: Presumably involved in the processing and regular
 CC turnover of intracellular proteins. Catalyzes the removal of
 CC unsubstituted N-terminal amino acids from various peptides (By
 CC similarity).
 CC -!- CATALYTIC ACTIVITY: Release of an N-terminal amino acid, Xaa-|-
 CC Xbb-, in which Xaa is preferably Leu, but may be other amino acids
 CC including Pro although not Arg or Lys, and Xbb may be Pro.
 CC -!- COFACTOR: Binds 2 manganese ions per subunit (By similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- SIMILARITY: Belongs to peptidase family M17.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF012163; AAM39965.1; -.
 DR HAMAP; MF_00181; -; 1.
 DR InterPro; IPR000819; Peptidase M17 C.
 DR InterPro; IPR008283; Peptidase_M17_N.
 DR Pfam; PF00883; Peptidase_M17; 1.
 DR Pfam; PF02789; Peptidase_M17_N; 1.
 DR PRINTS; PR00481; LAMNOPPTDASE.
 DR PROSITE; PS00631; CYTOSOL AP; 1.
 KW Hydrolase; Aminopeptidase; Manganese; Complete proteome.
 FT ACT_SITE 274 274 POTENTIAL.
 FT ACT_SITE 348 348 POTENTIAL.
 FT METAL 262 262 MANGANESE 2 (BY SIMILARITY).
 FT METAL 267 267 MANGANESE 1 AND 2 (BY SIMILARITY).
 FT METAL 285 285 MANGANESE 2 (BY SIMILARITY).
 FT METAL 344 344 MANGANESE 1 (BY SIMILARITY).
 FT METAL 346 346 MANGANESE 1 AND 2 (BY SIMILARITY).
 SQ SEQUENCE 493 AA; 51590 MW; 764A6FAD2BDB775 CRC64;
 Query Match 80.5%; Score 33; DB 1; Length 493;
 Best Local Similarity 85.7%; Pred. No. 27;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GNNYRPS 7
 DB 321 GNAYRPS 327
 RESULT 13
 RAG2_XENLA STANDARD; PRT; 520 AA.
 ID RAG2_XENLA
 AC Q91830;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE V(D)J recombination activating protein 2 (RAG-2).
 GN RAG2.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HD-1;
 RX MEDLINE=93389137; PubMed=8376769;
 RA Greenhalgh P.H., Olesen C.E., Steiner L.A.;
 RT "Characterization and expression of recombination activating genes
 RT (RAG-1 and RAG-2) in Xenopus laevis.";
 RL J. Immunol. 151:3100-3110(1993).

CC -!- FUNCTION: During lymphocyte development, the genes encoding
 CC immunoglobulins and T cell receptors are assembled from variable
 CC (V), diversity (D), and joining (J) gene segments. This
 CC combinatorial process, known as V(D)J recombination, allows the
 CC generation of an enormous range of binding specificities from a
 CC limited amount of genetic information. The RAG1/RAG2 complex
 CC initiates this process by binding to the conserved recombination
 CC signal sequences (RSS) and introducing a double-strand break
 CC between the RSS and the adjacent coding segment. These breaks are
 CC generated in two steps, nicking of one strand (hydrolysis),
 CC followed by hairpin formation (transesterification). RAG1/2 has
 CC also been shown to function as a transposase in vitro, and to
 CC possess RSS-independent endonuclease activity (end processing) and
 CC hairpin opening. RAG1 alone can bind to RSS but stable, efficient
 CC binding requires RAG2. All known catalytic activities require the
 CC presence of both proteins (By similarity).
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- TISSUE SPECIFICITY: Is expressed within the thymus, liver and
 CC spleen in juvenile frogs, and within the thymus and bone marrow of
 CC adults. A lower level expression is seen in the ovaries.
 CC -!- SIMILARITY: Belongs to the RAG2 family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L19325; AAA49943.1; -.
 DR PIR; I51556; I51556.
 DR InterPro; IPR004321; RAG2.
 DR Pfam; PF03089; RAG2; 1.
 KW Hydrolase; Endonuclease; Nuclear protein; DNA-binding;
 KW DNA recombination.
 SQ SEQUENCE 520 AA; 58636 MW; E5105425D52959DE CRC64;
 Query Match 80.5%; Score 33; DB 1; Length 520;
 Best Local Similarity 83.3%; Pred. No. 29;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GNNYRP 6
 DB 225 GNNFRP 230
 RESULT 14
 HPCL_MOUSE STANDARD; PRT; 581 AA.
 ID HPCL_MOUSE
 AC Q9QXEO; Q9DAV1;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE 2-hydroxyphytanoyl-CoA lyase (EC 4.1.-.-) (2-HPCL).
 GN HPCL.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Foulon V., Casteels M., Mannaerts G.P., Van Veldhoven P.P.;
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Placenta;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,

RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Moffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shermen C.M., Schuler G.D.,
RA Altschul S.F., Zebberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marushina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toehiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Munz D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J.J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinaki M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Catalyzes a carbon-carbon cleavage reaction; cleaves a
CC 2-hydroxy-3-methylacyl-CoA into formyl-CoA and a 2-methyl-branched
CC fatty aldehyde (By similarity).
CC -!- CATALYTIC ACTIVITY: 2-hydroxyphytanoyl-CoA = pristanal + formyl-
CC CoA.
CC -!- COFACTOR: Binds 1 thiamine pyrophosphate and 1 magnesium ion per
CC subunit (By similarity).
CC -!- PATHWAY: Alpha-oxidation of 3-methyl branched fatty acids
CC (phytanic acid); third step.
CC -!- SUBUNIT: Homotetramer (By similarity).
CC -!- SUBCELLULAR LOCATION: Peroxisomal (By similarity).
CC -!- SIMILARITY: Belongs to the TPP enzyme family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AJ132139; CAB65550.1; -;
CC DR EMBL: AK005505; BAB24085.1; -;
CC DR EMBL: BC021360; AAH21360.1; -;
CC DR MGD: MGI:1929657; Hpcl.
CC DR GO: GO:0016829; Fatty acid metabolism; IDA.
CC DR GO: GO:0006629; Lipid metabolism; IDA.
CC DR InterPro: IPR000399; Pyruvate decarb.
CC DR Pfam: PF00205; TPP enzymes; 1.
CC DR Pfam: PF02775; TPP enzymes; 1.
CC DR Pfam: PF02776; TPP enzymes; 1.
CC DR PROSITE: PS00187; TPP ENZYMES; FALSE NEG.
CC DR Flaprotin; Lyase; Thiamine pyrophosphate; Magnesium; Peroxisome.
CC -!- FUNCTION: Cadherins are calcium dependent cell adhesion proteins.
CC -!- They preferentially interact with themselves in a homophilic
CC SIMILARITY).

FT CONFLICT 252 252 M -> I (IN REF. 1).
SQ SEQUENCE 581 AA; 53660 MW; 7CABBFC780A32A69 CRC64;
Query Match 80.5%; Score 33; DB 1; Length 581;
Best Local Similarity 85.7%; Pred. No. 32;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GNNYRPS 7
DB 315 GNNVRPS 321
RESULT 15
CAQO_HUMAN STANDARD; PRT; 847 AA.
AC Q8IXH8; Q8TCH3; Q9BQ4; Q9NRU1;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cadherin-like protein 26 precursor (Cadherin-like protein VR20).
GN CDH26.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A. (ISOFORMS 2 AND 4).
RP TISSUE=lung, and Tonsillar carcinoma;
RA Nakajima T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K.,
RA Nakajima Y., Mizuno T., Morinaga M., Tanigami A., Fujiwara T., Ono T.,
RA Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y., Ota T., Suzuki Y.,
RA Oyayashi M., Nishi T., Shibahara T., Tanaka T., Nakamura Y.,
RA Isozaki T., Sugano S.;
RA Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RN SEQUENCE FROM N.A.
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baggeley C.L.,
RA Bailey J., Barlow K.P., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark A.P., Clark L.N., Clark S.Y., Clee C.M.,
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Graham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Leharvalho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McConachie L.J., McKay K., McMurray A.A.,
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.M., Ross M.T., Scott C.E., Senra H.K., Shownkeen R., Sims S.,
RA Skann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.;
RT "The DNA sequence and comparative analysis of human chromosome 20.";
RL Nature 414:865-871(2001).
RN [3]
RN SEQUENCE OF 405-847 FROM N.A. (ISOFORM 3).
RA Koola P.F.J., Van Roy F.;
RT "Identification and characterization of a novel human cadherin with
RT similarity to N-cadherin.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Cadherins are calcium dependent cell adhesion proteins.
CC -!- They preferentially interact with themselves in a homophilic

manner in connecting cells; cadherins may thus contribute to the sorting of heterogeneous cell types.

-!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).

-!- ALTERNATIVE PRODUCTS:

Event=Alternative splicing; Named isoforms=4;

Name=1 Synonyms=a;

isoId=Q81XH8-1; Sequence=Displayed;

Note=No experimental confirmation available;

Name=2; Synonyms=b;

isoId=Q81XH8-2; Sequence=VSP_008331, VSP_008333, VSP_008334;

Note=No experimental confirmation available;

Name=3;

isoId=Q81XH8-3; Sequence=VSP_008332;

Note=No experimental confirmation available;

Name=4;

isoId=Q81XH8-4; Sequence=VSP_008331;

Note=No experimental confirmation available;

-!- SIMILARITY: Contains 4 cadherin domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AK055202; -- NOT ANNOTATED_CDS.

DR EMBL; AK074477; BAB8093.1; --

DR EMBL; ALI09928; CAD5557.1; ALT_TERM.

DR EMBL; AF169690; AAF9687.1; --

DR EMBL; ALI09928; CAD5556.1; --

DR Genew; HGNC:15902; CDH26.

DR InterPro; IPR002126; Cadherin.

DR Pfam; PF00028; cadherin; 4.

DR PRINTS; PRO0205; CADHERIN.

DR SMART; SM00112; CA; 4.

DR PROSITE; PS00232; CADHERIN_1; 2.

DR PROSITE; PS0268; CADHERIN_2; 4.

KW Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;

KW Signal; Alternative splicing. POTENTIAL.

FT SIGNAL 1 26

FT CHAIN 27 847

FT DOMAIN 27 609

FT TRANSMEM 610 630

FT DOMAIN 631 847

FT DOMAIN 35 160

FT DOMAIN 161 270

FT DOMAIN 271 391

FT DOMAIN 392 495

FT CARBOHYD 76 76

FT CARBOHYD 80 80

FT CARBOHYD 166 166

FT CARBOHYD 172 172

FT CARBOHYD 457 457

FT VARSPLIC 694 694

FT VARSPLIC 720 847

FT VARSPLIC 720 847

FT KERNRFLSRLGCIIFQGRATAGRLPQDIYKEMPRRLTQ

FT TGRKRGALARTSPFKKVVYDHKEENKAGRKQRSHLFKV

FT MLQRNQGGRVQSASPSPLNKKACFPDGYRGSAGHNC

FT LARVSG -> KLHVANVLEDDPGYLPHVYSEGECCGAPSL

FT SLASLQELQDLDLSLGSKATPFEEIYESGVPS (in isoform 3).

FT /FTid=VSP_008332.

FT Missing (in isoform 2).

FT /FTid=VSP_008333.

FT KGTSAQ -> MKPLIW (in isoform 2).

FT /FTid=VSP_008334.

FT 94615 MW; 48A378DC65SEB5D CRC64;

FT SEQUENCE 847 AA; 94615 MW; 48A378DC65SEB5D CRC64;

Query Match 80.5%; Score 33; DB 1; Length 847;

Best Local Similarity 71.4%; Pred. No. 47;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNYRPS 7

DB 263 GNNHRPA 269

RESULT 16

MSRA_STRP3 STANDARD; PRT; 169 AA.

AC Q8KSE4;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Peptide methionine sulfoxide reductase msrA (EC 1.8.4.6) (Protein-methionine-S-oxide reductase) (Peptide Met(O) reductase).

GN MSRA.2 OR SPYW3_0329 OR SP81528. M3.

OS Streptococcus pyogenes (serotype M3).

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

OC Streptococcus.

OX NCBI_TaxID=198466;

RN [1]_TaxID=198466;

RP SEQUENCE FROM N.A.

RC STRAIN=MGAS315 / Serotype M3;

RX MEDLINE=22133808; PubMed=12122206;

RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D., Beres S.B., Sylva G.L.,

RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M., Schlievert P.M., Musser J.M.;

RT "Genome sequence of a serotype M3 strain of group A Streptococcus: phage-encoded toxins, the high-virulence phenotype, and clone emergence.";

RT Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083 (2002).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=SSI-1 / Serotype M3;

RX MEDLINE=22683278; PubMed=12799345;

RA Nakagawa I., Kurokawa K., Yamashita A., Nakata M., Tomiyasu Y., Okahashi N., Kawabata S., Yamazaki K., Shiba T., Yasunaga T., Hayashi H., Hattori M., Hamada S.;

RT "Genome sequence of an M3 strain of Streptococcus pyogenes reveals a large-scale genomic rearrangement in invasive strains and new insights into phage evolution.";

RT Genome Res. 13:1042-1055 (2003).

RL

CC -!- FUNCTION: Has an important function as a repair enzyme for proteins that have been inactivated by oxidation. Catalyzes the reversible oxidation-reduction of methionine sulfoxide in proteins to methionine (By similarity).

CC -!- CATALYTIC ACTIVITY: Protein L-methionine + oxidized thioredoxin = protein L-methionine S-oxide + reduced thioredoxin.

CC -!- SIMILARITY: Belongs to the msrA Met sulfoxide reductase family.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC

CC EMBL; AE014143; AAM78936.1; --

CC EMBL; AP005145; BAC64623.1; --

DR HAMAP; MF 01401; --; 1.

DR InterPro; IPR002569; PMSR.

DR Pfam; PF01625; PMSR; 1.

DR ProDom; PD003489; PMSR; 1.

DR TIGRFAMs; TIGR00401; msrA; 1.

KW Oxidoreductase; Complete proteome.

FT ACT_SITE 10 10

FT BY SIMILARITY.

FT SEQUENCE 169 AA; 19512 MW; 66E02871CC279F7B CRC64;

Query Match 78.0%; Score 32; DB 1; Length 169;

Best Local Similarity 83.3%; Pred. No. 14;

```

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNYRP 6
Db 90 GDNYP 95

RESULT 17
MSRA_STRP8 STANDARD; PRT; 169 AA.
AC Q8P272;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Peptide methionine sulfoxide reductase msra (EC 1.8.4.6) (Protein-
GN MSRA OR SPY18_0509.
OS Streptococcus pyogenes (serotype M18).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=186103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MGAS8232 / Serotype M18;
RX MEDLINE=21927593; PubMed=11917108;
RA Smoot J.C., Barbican K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
RA Sylva G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,
RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
RT "Genome sequence and comparative microarray analysis of serotype M18
RT group A Streptococcus strains associated with acute rheumatic fever
RT outbreaks.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673 (2002).
CC -1- FUNCTION: Has an important function as a repair enzyme for
CC proteins that have been inactivated by oxidation. Catalyzes the
CC reversible oxidation-reduction of methionine sulfoxide in proteins
CC to methionine (By similarity).
CC -1- CATALYTIC ACTIVITY: Protein L-methionine + oxidized thioredoxin =
CC protein L-methionine S-oxide + reduced thioredoxin.
CC -1- SIMILARITY: Belongs to the msra Met sulfoxide reductase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB009989; AAL97226.1; -.
DR HAMAP; MF_01401; -.
DR InterPro; IPR002569; PMSR.
DR Pfam; PF01625; PMSR; 1.
DR ProDom; PD003489; PMSR; 1.
DR TIGRFAMs; TIGR00401; msra; 1.
KW Oxidoreductase; Complete proteome.
FT ACT_SITE 10 10 BY SIMILARITY.
SQ SEQUENCE 169 AA; 15483 MW; 7800C67F3237891C CRC64;

Query Match 78.0%; Score 32; DB 1; Length 169;
Best Local Similarity 83.3%; Pred. No. 14;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNYRP 6
Db 90 GDNYP 95

RESULT 18
MSRA_STRP STANDARD; PRT; 169 AA.
AC Q9A139;
DT 28-FEB-2003 (Rel. 41, Created)
DE Peptide methionine sulfoxide reductase msra 2 (EC 1.8.4.6) (Protein-
GN MSRA2 OR PMSR OR LL2006.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IL1403;
RX MEDLINE=21235186; PubMed=11337471;
RA Bolotin A., Wincker P., Manger S., Jaillon O., Malmgren K.,

```

```

DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Peptide methionine sulfoxide reductase msra (EC 1.8.4.6) (Protein-
DE methionine-S-oxide reductase) (Peptide Met(O) reductase).
GN MSRA OR MSRA.2 OR SPY0466.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SF370 / ATCC 700294 / Serotype M1;
RX MEDLINE=21192684; PubMed=11296296;
RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663 (2001).
CC -1- FUNCTION: Has an important function as a repair enzyme for
CC proteins that have been inactivated by oxidation. Catalyzes the
CC reversible oxidation-reduction of methionine sulfoxide in proteins
CC to methionine (By similarity).
CC -1- CATALYTIC ACTIVITY: Protein L-methionine + oxidized thioredoxin =
CC protein L-methionine S-oxide + reduced thioredoxin.
CC -1- SIMILARITY: Belongs to the msra Met sulfoxide reductase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB006506; AAK33479.1; -.
DR HSSP; P54149; 1FVG.
DR HAMAP; MF_01401; -.
DR InterPro; IPR002569; PMSR.
DR Pfam; PF01625; PMSR; 1.
DR ProDom; PD003489; PMSR; 1.
DR TIGRFAMs; TIGR00401; msra; 1.
KW Oxidoreductase; Complete proteome.
FT ACT_SITE 10 10 BY SIMILARITY.
SQ SEQUENCE 169 AA; 15482 MW; 5EE1E1823A1BE30 CRC64;

Query Match 78.0%; Score 32; DB 1; Length 169;
Best Local Similarity 83.3%; Pred. No. 14;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNYRP 6
Db 90 GDNYP 95

RESULT 19
MSA2_LACLA STANDARD; PRT; 172 AA.
AC Q9CE42;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Peptide methionine sulfoxide reductase msra 2 (EC 1.8.4.6) (Protein-
DE methionine-S-oxide reductase 2) (Peptide Met(O) reductase 2).
GN MSRA2 OR PMSR OR LL2006.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IL1403;
RX MEDLINE=21235186; PubMed=11337471;
RA Bolotin A., Wincker P., Manger S., Jaillon O., Malmgren K.,

```

RA Weissenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium *Lactococcus*
RT *lactis* ssp. *lactis* I11403.";
RL Genome Res. 11:731-753(2001).
CC -!- FUNCTION: Has an important function as a repair enzyme for
CC proteins that have been inactivated by oxidation. Catalyzes the
CC reversible oxidation-reduction of methionine sulfoxide in proteins
CC to methionine (By similarity).
CC -!- CATALYTIC ACTIVITY: Protein L-methionine + oxidized thioredoxin =
CC protein L-methionine S-oxide + reduced thioredoxin.
CC -!- SIMILARITY: Belongs to the msrA Met sulfoxide reductase family.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; A5006430; AAK06104.1; -;
DR HSSP; P54149; 1FVG.
DR HAMAP; MF_01401; -; 1.
DR InterPro; IPR002569; PMSR.
DR Pfam; PF01625; PMSR; 1.
DR ProDom; PD003489; PMSR; 1.
DR TIGRFAMs; TIGR00401; msrA; 1.
KW Oxidoreductase; Complete proteome.
KW ACT_SITE 12 12 BY SIMILARITY.
FT ACT_SITE 12 12 BY SIMILARITY.
SQ SEQUENCE 172 AA; 19878 MW; B84CCDC5FE7F06F CRC64;

Query Match 78.0%; Score 32; DB 1; Length 172;
Best Local Similarity 83.3%; Pred. No. 15;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNYRP 6
Db 92 GDNYRP 97
|:|||||

RESULT 20
THPS_SULAC STANDARD; PRT; 340 AA.
ID THPS_SULAC STANDARD; PRT; 340 AA.
AC P17118;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Thermoplasma precursor (EC 3.4.23.42).
OS Sulfolobus acidocaldarius.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus
OX NCBI_TaxID=2285;
RN [1]
SEQUENCE FROM N.A., AND SEQUENCE OF 42-76.
RP MEDLINE=90110209; PubMed=2104843;
RX Lin X.-L., Tang J.;
RA "Purification, characterization, and gene cloning of thermoplasma, a
RT thermotable acid protease from Sulfolobus acidocaldarius.";
RL J. Biol. Chem. 265:1490-1495(1990).
CC -!- FUNCTION: MAY REPRESENT A NEW CLASS OF ACID PROTEASES. IT DIGEST
CC PROTEINS AND PEPTIDES IN ACIDIC SOLUTION, AND IS THERMOSTABLE. IT
CC HAS MAXIMAL PROTEOLYTIC ACTIVITY AT PH 2 AND 90 DEGREES CELSIUS.
CC -!- CATALYTIC ACTIVITY: Specificity similar to pepsin A, prefers bulky
CC hydrophobic side-chains on either side of the scissile bond.
CC -!- SUBCELLULAR LOCATION: MAY BE LINKED TO CELLS BY COVALENT LINKAGES
CC THROUGH SOME SIDE CHAINS.
CC -!- MISCELLANEOUS: THERMOPLASMA HAS PERHAPS A DIFFERENT TYPE OF ACTIVE
CC SITE THAN OTHER ASPARTIC PROTEASES.
CC -!- SIMILARITY: Belongs to peptidase family A5.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; J05184; AAA72221.1; -;
DR MEROPS; A05.001; -;
DR InterPro; IPR007981; Peptidase_A5.
DR Pfam; PF05317; Thermoplasin; 1. Signal; Glycoprotein.
KW Hydrolase; Aspartyl protease; POTENTIAL.
FT SIGNAL 1 28
FT CHAIN 29 41 SHORT (POTENTIAL).
FT CARBOHYD 31 31 N-LINKED (POTENTIAL).
FT CARBOHYD 65 65 N-LINKED (PROBABLE).
FT CARBOHYD 69 69 N-LINKED (PROBABLE).
FT CARBOHYD 85 85 N-LINKED (POTENTIAL).
FT CARBOHYD 117 117 N-LINKED (POTENTIAL).
FT CARBOHYD 148 148 N-LINKED (POTENTIAL).
FT CARBOHYD 197 197 N-LINKED (POTENTIAL).
FT CARBOHYD 277 277 N-LINKED (POTENTIAL).
FT CARBOHYD 287 287 N-LINKED (POTENTIAL).
FT CARBOHYD 327 327 N-LINKED (POTENTIAL).
FT CARBOHYD 334 334 N-LINKED (POTENTIAL).
SQ SEQUENCE 340 AA; 37262 MW; E2758C0BC287FCE1 CRC64;

Query Match 78.0%; Score 32; DB 1; Length 340;
Best Local Similarity 71.4%; Pred. No. 30;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
Db 246 GNNYTPN 252
|:|||||

RESULT 21
ARPA_HUMAN STANDARD; PRT; 417 AA.
ID ARPA_HUMAN STANDARD; PRT; 417 AA.
AC Q9NZ32; Q9H9Y5; Q9NWY2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Actin-related protein 10 (hARP11).
GN ACTR10 OR ACTR11 OR ARP11.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RP TISSUE=Hypothalamus;
RC Xiao H., Song H., Gao G., Ren S., Chen Z., Han Z.;
RA "A novel gene expressed in human hypothalamus.";
RT Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RL [2]
SEQUENCE FROM N.A.
RN [3]
RA Isoqai T., Ota T., Hayaishi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA Wagatsuma M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
RA Ninomiya K., Iwayanagi T.;
RT "NEO human cDNA sequencing project.";
RN Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RL [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=2388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner C.M., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muray D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Whiting J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- SIMILARITY: Belongs to the actin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF220190; AAF67655.1; -;
DR EMBL; AK000544; BAA91243.1; -;
DR EMBL; AK022534; BAB14083.1; -;
DR EMBL; BC011997; AAH11997.1; -;
DR Genew; HGNC:17372; ACTR10.
DR InterPro; IPR004000; Actin_like.
DR SMART; SM00268; ACTIN; 1.
DR Structural protein; Cytoskeleton.
FT CONFLICT 53 53 R -> K (IN REF. 2; BAA91243).
FT CONFLICT 93 93 E -> K (IN REF. 2; BAA91243).
FT CONFLICT 219 219 D -> G (IN REF. 2; BAB14083).
SQ SEQUENCE 417 AA; 46307 MW; DF351A7B299DB7C1 CRC64;

Query Match 78.0%; Score 32; DB 1; Length 417;
Best Local Similarity 85.7%; Pred. No. 37;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
DB 235 GNNRPS 241

RESULT 22
PO23 RAT STANDARD; PRT; 430 AA.
AC P42571; P42572; PRT;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE POU domain, class 2, transcription factor 3 (Octamer-binding
DE transcription factor 11) (Oct-11) (Transcription factor Skn-1).
GN POU2F3 OR OTF11 OR SKN1 OR SKN-1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Fischer; TISSUE=Skin;
RX MEDLINE=93219836; PubMed=7682011;
RA Andersen B., Schonemann M.D., Flynn S.E., Pearse R.V. II,
RA Singh H., Rosenfeld M.G.;
RT "Skn-1a and Skn-1i: two functionally distinct Oct-2-related factors
RT expressed in epidermis.";

RL Science 260:78-82(1993).
RN [2]
RX ERRATUM.
RP MEDLINE=94069332; PubMed=8248794;
RA Andersen B., Schonemann M.D., Flynn S.E., Pearse R.V. II,
RA Singh H., Rosenfeld M.G.;
RA Science 262:1499-1495(1993).
RL -!- FUNCTION: Transcription factor that binds to the octamer motif
CC ('ATTTCAT'). Activates cyokeratin 10 (K10) gene expression. May
CC serve a regulatory function with respect to epidermal development.
CC Isoform 2 inhibits transactivation by OCT-1.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1; Synonyms=Skn-1a;
CC IsoID=P42571-1; Sequences=Displayed;
CC Name=2; Synonyms=Skn-1i;
CC IsoID=P42571-2; Sequences=VSP 002332;
CC -!- TISSUE SPECIFICITY: Expressed in epidermis.
CC -!- SIMILARITY: Belongs to the POU transcription factor family.
CC -!- SIMILARITY: Contains 1 homeobox domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L23862; -; NOT_ANNOTATED_CDS.
DR EMBL; L23863; -; NOT_ANNOTATED_CDS.
DR HSSP; FL4859; 1OCT.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR000972; Octamer-bind_TF.
DR InterPro; IPR000327; POU domain.
DR InterPro; IPR007103; POU_homeo.
DR Pfam; PF00046; homeobox; 1.
DR Pfam; PF00157; pou; 1.
DR PRINTS; PR00029; OCTAMER.
DR PRINTS; PR00028; POUDOMAIN.
DR ProDom; PD000010; Homeobox; 1.
DR ProDom; PD000583; POU domain; 1.
DR SMART; SM00389; HOX; 1.
DR SMART; SM00352; POU; 1.
DR PROSITE; PS00035; POU 1; 1.
DR PROSITE; PS00465; POU 2; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
KW Homeobox; DNA-binding; transcription regulation; Activator;
KW Nuclear protein; Alternative splicing.
FT DOMAIN 180 250
FT DNA_BIND 274 333
FT DOMAIN 347 419
FT VARSPLIC 1 113
SQ SEQUENCE 430 AA; 46813 MW; 9B538C8D9743B96F CRC64;

Query Match 78.0%; Score 32; DB 1; Length 430;
Best Local Similarity 85.7%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
DB 379 GNNRPS 385

RESULT 23
PO23 MOUSE

ID PO23_MOUSE STANDARD; PRT; 431 AA.
 AC F31362;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE POU domain, class 2, transcription factor 3 (Octamer-binding
 transcription factor 1) (Oct-1) (Epoc-1).
 GN POU2F3 OR OTF11 OR OTF11 OR OTF11 OR EPOC1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Thymus;
 RX MEDLINE=93181164; PubMed=8441607;
 RA Goldsborough A.S., Healy L.E., Copeland N.G., Gilbert D.J.,
 RA Jenkins N.A., Willison K.R., Ashworth A.;
 RT "Cloning, chromosomal localization and expression pattern of the POU
 domain gene Oct-11.";
 RL Nucleic Acids Res. 21:127-134(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/c; TISSUE=Thymus;
 RX MEDLINE=94040806; PubMed=8224904;
 RA Yukawa K., Yasui T., Yamamoto A., Shiku H., Kishimoto T.,
 RA Kikutani H.;
 RT "Epoc-1: a POU-domain gene expressed in murine epidermal basal cells
 and thymic stromal cells.";
 RL Gene 133:163-169(1993).
 CC -!- FUNCTION: Transcription factor that binds to the octamer motif
 ('ATTGTCAT').
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- TISSUE SPECIFICITY: Skin, thymus, stomach and testis.
 CC -!- DEVELOPMENTAL STAGE: During embryogenesis and in adults.
 CC -!- SIMILARITY: Belongs to the POU transcription factor family.
 CC Class-2 subfamily.
 CC -!- SIMILARITY: Contains 1 homeobox domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 DR EMBL; Z18537; CAA79222.1; -.
 DR EMBL; L14677; AAA16855.1; -.
 DR PIR; S35541; S35541.
 DR HSP; P14859; IOCT.
 DR TRANSFAC; T01874; -.
 DR MGD; MGI:102585; Pou2f3.
 DR InterPro; IPR001356; Homeobox.
 DR InterPro; IPR000972; Octamer-bind_TF.
 DR InterPro; IPR000327; POU domain.
 DR InterPro; IPR007103; POU homeo.
 DR Pfam; PF00046; homeobox_1.
 DR Pfam; PF00157; pou; 1.
 DR PRINTS; PR00029; OCTAMER.
 DR PRINTS; PR00028; POUDOMAIN.
 DR ProDom; PD000010; Homeobox; 1.
 DR ProDom; PD000583; POU domain; 1.
 DR SMART; SM00389; HOX; 1.
 DR SMART; SM00352; POU; 1.
 DR PROSITE; PS00027; HOMEBOX_1; 1.
 DR PROSITE; PS00035; POU 1; 1.
 DR PROSITE; PS00465; POU 2; 1.
 DR PROSITE; PS50071; HOMEBOX_2; 1.
 KW DNA-binding; Nuclear protein; Transcription regulation; Homeobox.
 FT DOMAIN 180 250
 FT DNA_BIND 274 333 HOMEBOX.
 FT DOMAIN 347 420 SER-RICH.

FT CONFLICT 139 139 R -> A (IN REF. 1).
 FT CONFLICT 249 249 P -> A (IN REF. 1).
 FT CONFLICT 396 431 PTASNNKAAANNSSSSSSSSSGSWYRNHPYLH ->
 FT SOVWALLT (IN REF. 1).
 SQ SEQUENCE 431 AA; 47071 MW; 7E283E797EA8D3FA CRC64;
 Query Match 78.0%; Score 32; DB 1; Length 431;
 Best Local Similarity 85.7%; Pred. No. 38;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GNNYRPS 7
 DB 379 GNNSRPS 385
 RESULT 24
 ID PO23_HUMAN STANDARD; PRT; 436 AA.
 AC Q9UKI9; Q9UKR7; Q9Y504;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE POU domain, class 2, transcription factor 3 (Octamer-binding
 transcription factor 1) (Oct-1) (Transcription factor Skn-1)
 DE (PLA-1 protein).
 DE POU2F3 OR OTF11 OR PLA1.
 OS Homo sapiens (Human).
 GN Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99403087; PubMed=10473598;
 RA Hildesheim J., Foster R.A., Chamberlin M.E., Vogel J.C.;
 RT "Characterization of the regulatory domains of the human skn-1a/Epoc-
 1/Oct-1 POU transcription factor.";
 RL J. Biol. Chem. 274:26399-26406(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Fischer D.F., Hemelaar J., Backendorf C.;
 RT "Characterisation of the human transcription factor Oct-11:
 involvement in the regulation of the SPRR2A gene.";
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Transcription factor that binds to the octamer motif
 ('ATTGTCAT'). Regulated the expression of a number of genes such
 as SPRR2A or placental lactogen.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- TISSUE SPECIFICITY: Specifically expressed in epidermis and
 cultured keratinocytes.
 CC -!- SIMILARITY: Belongs to the POU transcription factor family.
 CC Class-2 subfamily.
 CC -!- SIMILARITY: Contains 1 homeobox domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 DR EMBL; AF162278; AAF00199.1; -.
 DR EMBL; AF133895; AAD55245.1; -.
 DR EMBL; AJ012214; CAB45383.1; -.
 DR HSP; P14859; IOCT.
 DR Genew; HGNC:19864; POU2F3.

```
DR MIM: 607394;
DR InterPro: IPR001356; Homeobox.
DR InterPro: IPR000972; Octamer-bind TF.
DR InterPro: IPR000327; POU domain.
DR InterPro: IPR007103; POU homeo.
DR Pfam: PF00046; homeobox; 1.
DR Pfam: PF00157; POU; 1.
DR PRINTS: PR00029; OCTAMER.
DR PRINTS: PR00028; POU DOMAIN.
DR ProDom: PD000010; Homeobox; 1.
DR ProDom: PD000583; POU domain; 1.
DR SMART: SM00389; HOX; 1.
DR SMART: SM00352; POU; 1.
DR PROSITE: PS00027; HOMEBOX 1; 1.
DR PROSITE: PS00071; HOMEBOX 2; 1.
DR PROSITE: PS00035; POU 1; 1.
DR PROSITE: PS00465; POU 2; 1.
KW DNA-binding; Nuclear protein; Transcription regulation; Homeobox.
FT DOMAIN 183 257
FT DNA_BIND 281 340
FT DOMAIN 354 425
FT CONFLICT 6 6
FT CONFLICT 37 37
FT CONFLICT 152 152
FT CONFLICT 351 351
SQ SEQUENCE 436 AA; 47451 MW; C674482739963C5B CRC64;

Query Match 78.0%; Score 32; DB 1; Length 436;
Best Local Similarity 85.7%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
DB 386 GNNRSRPS 392

RESULT 25
Z224_HUMAN
ID Z224_HUMAN STANDARD; PRT; 707 AA.
AC Q9NZL3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE Zinc finger protein 224.
GN ZNF224.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Shannon M., Branscomb E., Hauser L., Gordon L., Ashworth L.,
RA Stubbs L.;
RT "Differential expansion of homologous zinc-finger gene families in
RT human chromosome 19q13.2 and mouse chromosome 7.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: May function as a transcription factor.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
CC FINGER PROTEINS.
CC -!- SIMILARITY: Contains 1 KRAB domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AF187990; AAF04106.2; -.
CC HSP: P08047; ISP2.
CC Genew: HGNC:13017; ZNF224.
DR
```

```
DR InterPro: IPR001909; KRAB.
DR InterPro: IPR007087; Znf_C2H2.
DR Pfam: PF01352; KRAB; 1.
DR Pfam: PF00096; zf-C2H2; 18.
DR ProDom: PD000003; Znf_C2H2; 9.
DR SMART: SM00349; KRAB; 1.
DR SMART: SM00355; Znf_C2H2; 18.
DR PROSITE: PS08005; KRAB; 1.
DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 17.
DR PROSITE: PS0157; ZINC_FINGER_C2H2_2; 18.
KW Transcription regulation; DNA-binding; Zinc-finger; Metal-binding;
KW Nuclear protein; Repeat.
FT DOMAIN 8 78
FT ZN_FING 176 198
FT ZN_FING 204 226
FT ZN_FING 232 254
FT ZN_FING 260 282
FT ZN_FING 288 310
FT ZN_FING 316 338
FT ZN_FING 344 366
FT ZN_FING 372 394
FT ZN_FING 400 422
FT ZN_FING 428 450
FT ZN_FING 456 478
FT ZN_FING 484 506
FT ZN_FING 512 534
FT ZN_FING 568 590
FT ZN_FING 596 618
FT ZN_FING 652 674
FT ZN_FING 680 702
SQ SEQUENCE 707 AA; 82267 MW; 157D9F7D4FCE39A7 CRC64;

Query Match 78.0%; Score 32; DB 1; Length 707;
Best Local Similarity 71.4%; Pred. No. 63;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
DB 150 GNGYRPS 156

Search completed: September 24, 2004, 01:55:17
Job time: 9.81132 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 01:47:17 ; Search time 33.5472 Seconds
(without alignments)
65.836 Million cell updates/sec

Title: US-10-088-639A-2_COPY_49_55
Perfect score: 41
Sequence: 1 GNNYRPS 7

Scoring table! BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database : SPTREMBL_25:*
1: sp archaea:*
2: sp bacteria:*
3: sp fungi:*
4: sp human:*
5: sp invertebrate:*
6: sp mammal:*
7: sp mhc:*
8: sp organelle:*
9: sp phase:*
10: sp plant:*
11: sp rodent:*
12: sp virus:*
13: sp vertebrate:*
14: sp unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	37	90.2	146	10	Q84KK9
2	37	90.2	327	5	Q9W316
3	37	90.2	447	5	Q8GBA6
4	37	90.2	569	5	Q8WRF2
5	37	90.2	742	13	Q9DE61
6	37	90.2	866	5	Q9VHA6
7	37	90.2	907	13	Q9DFB2
8	36	87.8	551	4	Q9NS70
9	36	87.8	551	4	Q9GJ63
10	36	87.8	572	4	Q9NYQ4
11	36	87.8	585	4	Q9NYQ5
12	36	87.8	588	4	Q8WYA1
13	36	87.8	599	4	Q8WYA2
14	36	87.8	602	4	Q9H2M4
15	36	87.8	622	4	Q8WYA3
16	36	87.8	636	4	Q8WYA4

Q80VW7 mus musculus
Q81KH7 heliothis z
Q810d5 microsporium
Q8nib6 trichophyto
Q8pub4 methanosarc
Q87fj7 vibrio para
Q8ex31 mycoplasma
Q86240 plasmodium
Q7x3p2 staphylococ
Q7ws19 staphylococ
Q39500 cylindrothe
Q849w3 escherichia
Q64340 bacterioph
Q7y3y7 bacterioph
Q8e8l4 shewanella
Q8chm7 rattus norv
Q7ze96 human immun
Q8e434 streptococ
Q8dyh1 streptococ
Q8dsy4 streptococ
Q834h1 enterococu
Q88vg9 lactobacill
Q86sy7 homo sapien
Q86ty2 homo sapien
Q9bvs9 homo sapien
Q91ks8 arabidopsis
Q9p555 neurospora
Q7wze6 shewanella
Q9gpk1 heterodera
Q8f4q5 leptospira
Q8kjd9 rhizobium l
Q8kjd9 rhizobium l
Q9e513 loofah wite
Q8kjd5 rhizobium l
Q7xnr8 oryza sativ
Q7xns1 oryza sativ
Q919k7 pasteurella
Q9ahm6 pasteurella
Q816u6 craterostig
Q9a0q8 streptococ
Q9epc1 methanosarc
Q9zw78 arabidopsis
Q8wqm5 strongyloce
Q9c8g7 arabidopsis
Q98qe4 mycoplasma
Q81zc8 homo sapien
Q8n1h1 trichophyto
Q8e1m3 shewanella
Q7xqr4 oryza sativ
Q9rna8 microcystis
Q9ynw1 anabaena sp
Q8yzq2 anabaena sp
Q86v10 homo sapien
Q9zq08 arabidopsis
Q7tn23 mus musculus
Q9a144 arabidopsis
Q7vtx9 drosophila
Q9m6s0 populus tre
Q7c071 gallus gall
Q73875 brachydanio
Q7lqc6 mus musculus
Q804v4 xenopus lae
Q7zt67 xenopus lae
Q8kcs5 mus musculus
Q9eq30 mus musculus
Q9zt18 xenopus lae
Q86xc7 homo sapien
Q9nj55 anopheles g
Q9nat0 anopheles g
Q9p247 homo sapien
Q9vzs2 drosophila
Q917t7 drosophila
Q8tf72 homo sapien

```

90 32 78.0 2457 12 041965
91 31 75.6 32 11 Q8C2N8
92 31 75.6 33 2 Q93RM5
93 31 75.6 98 16 Q7V1C6
94 31 75.6 99 15 Q760S3
95 31 75.6 99 15 Q9YU06
96 31 75.6 99 15 Q9W9R3
97 31 75.6 99 15 Q9YU05
98 31 75.6 99 15 Q7SVK1
99 31 75.6 99 15 Q7SVK0
100 31 75.6 99 15 Q7SVJ9
101 31 75.6 99 15 Q7SVJ7
102 31 75.6 99 15 Q7SVJ6
103 31 75.6 99 15 Q7SVJ1
104 31 75.6 149 16 Q8NLG6
105 31 75.6 150 15 Q90066
106 31 75.6 184 16 P74110
107 31 75.6 212 5 Q9XWT4
108 31 75.6 234 16 Q8ZA25
109 31 75.6 236 16 Q8ZF15
110 31 75.6 250 2 Q9EZY2
111 31 75.6 250 2 Q9EZY1
112 31 75.6 251 12 Q9PYU2
113 31 75.6 253 2 Q93OB8
114 31 75.6 253 2 Q7XOM7
115 31 75.6 253 2 Q7XOM6
116 31 75.6 253 2 Q7XOM5
117 31 75.6 253 2 Q7WS24
118 31 75.6 253 2 Q7WS21
119 31 75.6 253 2 Q7WRH2
120 31 75.6 257 5 Q9NSD8
121 31 75.6 258 10 Q9ZTC2
122 31 75.6 263 2 Q93QA3
123 31 75.6 264 2 Q93OB7
124 31 75.6 265 2 Q93QA2
125 31 75.6 265 2 Q93QA5
126 31 75.6 265 2 Q93QB6
127 31 75.6 265 2 Q93QA4
128 31 75.6 266 2 Q93QB2
129 31 75.6 266 2 Q93QA9
130 31 75.6 267 12 Q7YT6
131 31 75.6 269 2 Q93OB0
132 31 75.6 272 2 Q93QC3
133 31 75.6 273 2 Q9EZW3
134 31 75.6 273 2 Q9EZW9
135 31 75.6 274 2 Q93QB9
136 31 75.6 274 2 Q93QA8
137 31 75.6 274 2 Q93C0
138 31 75.6 274 2 Q93QA7
139 31 75.6 274 2 Q93QA6
140 31 75.6 275 2 Q93QC2
141 31 75.6 275 2 Q93QB5
142 31 75.6 275 2 Q93QB3
143 31 75.6 275 2 Q9EZY3
144 31 75.6 275 2 Q93QC1
145 31 75.6 275 2 Q93QB4
146 31 75.6 275 2 Q9EZW5
147 31 75.6 275 2 Q9EZW4
148 31 75.6 275 2 Q93QB1
149 31 75.6 277 2 Q9EZX1
150 31 75.6 277 2 Q9EZX9

```

ALIGNMENTS

```

RESULT 1
Q84KK9 PRELIMINARY; PRT; 146 AA.
ID Q84KK9;
AC Q84KK9;
DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)

```

```

DE Jacalin-related lectin.
GN PALB.
OS Phlebotomus aureum.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Moniliformopses; Filicophyta; Filicopsida; Filicales; Polypodiaceae;
OC Phlebotomus.
OX NCBI_TaxID=218620;
RN [1]
SEQUENCE FROM N.A.
RC TISSUE=Rhizome;
RX MEDLINE=22538482; PubMed=12538584;
RA Tateno H., Harry W.C., Petryniak J., Goldstein I.J.;
RT "Purification, Characterization, Molecular Cloning, and Expression of
RT Novel Members of Jacalin-related Lectins from Rhizomes of the True
RT Fern Phlebotomus aureum (L.) J. Smith (Polypodiaceae).";
RL J. Biol. Chem. 278:10891-10899(2003).
DR EMBL; AB099933; BAC55269.1;
DR GO; GO:0005529; P:sugar binding; IEA.
DR GO; GO:0007157; P:heterophilic cell adhesion; IEA.
DR InterPro; IPR001229; Jacalin_lectin.
KW Lectin.
SQ SEQUENCE 146 AA; 15026 MW; 55353B28E1CEA6A CRC64;
Query Match 90.2%; Score 37; DB 10; Length 146;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GNNYRP 6
DB 117 GNNYRP 122
|||||
RESULT 2
Q9W316 PRELIMINARY; PRT; 327 AA.
ID Q9W316; Q961F5;
AC Q9W316; Q961F5;
DT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-OCT-2003 (T-EMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE CG32698 protein (GH26663p).
GN CG32698 OR CG9678 OR CG9683.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abriel J.F., Agbayani A., An H.-J., Andrews-Frannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Fabros B., Deicher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

```

RA Merkulo G., Milshina N.V., Moberly C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Celniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Amanatides J.G., Brandon R.C., Rogers Y.,
RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT "Sequencing of *Drosophila melanogaster* genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA Tupy J.L., Bergman C., Carlson J.W., Celniker S.E.,
RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA Krommiller B., Marshall B., Milburn G., Richter J., Russo S.,
RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA Ashburner M., Gelbart W.N., Rubin G.M., Mungall C.J., Lewis S.E.;
RT "Annotation of *Drosophila melanogaster* genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RA Stapleton M., Brokstein P., Hong L., Agbavani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF003448; AAF46521.2; -;
DR EMBL; AY051620; AAK93044.1; -;
DR HSSP; O43570; 1JDO.
DR FlyBase; FBN0052698; CG32698.
DR GO; GO:004089; F:carbonate dehydratase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006730; P:one-carbon compound metabolism; IEA.
DR InterPro; IPR001148; Euk COAnhd.
DR Pfam; PF00194; carb anhydrase; 1.
DR ProDom; PD000865; Euk_COAnhd; 2.
SQ SEQUENCE 327 AA; 37182 MW; 9574C08B8BA9FD01 CRC64;

Query Match

90.2%; Score 37; DB 5; Length 327;

Best Local Similarity 100.0%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GNNYRP 6
DB 279 GNNYRP 284

RESULT 3
Q86BA6 PRELIMINARY; PRT; 447 AA.
AC Q86BA6;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DE CG8312-PB.
GN CG8312
OS *Drosophila melanogaster* (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; *Drosophila*.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides J.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbavani A., An H.J., Andrews-Pfannkuch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Chertys K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Hostin D., Houston K.A., Howland T.J., Hernandez J.R., Houck J.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Moberly C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA Tupy J.L., Bergman C.M., Berman B.P., Carlson J.W., Celniker S.E.,
RA Clamp M.E., Drysdale R.A., Emmert D., Frise E., de Grey A.D.N.J.,
RA Harris N.L., Krommiller B., Marshall B., Milburn G.H., Richter J.,
RA Russo S., Searle S.M.J., Smith E., Shu S., Smutniak F.,
RA Whitfield E.J., Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J.,
RA Lewis S.E.;

RT "Annotation of Drosophila melanogaster genome.";
 RN Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RL [3]

RP SEQUENCE FROM N.A.
 RA FlyBase;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 RN [4]

RP SEQUENCE FROM N.A.
 RA FlyBase;
 RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AEO03683; AAO41527.1; -
 SQ SEQUENCE 447 AA; 48394 MW; 5AB7FDG29E1B62FE CRC64;

Query Match 90.2%; Score 37; DB 5; Length 447;
 Best Local Similarity 100.0%; Pred. No. 43;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNYRP 6
 Db 6 GNNYRP 11
 |||||

RESULT 4

Q8MRF2 ID Q8MRF2 PRELIMINARY; PRT; 569 AA.
 AC Q8MRF2;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE GH04568p.
 GN CG8312.

OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkeley;
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Friese E.,
 RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Celiker S.;
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY121618; AAMS1945.1; -
 DR FlyBase; FBgn0037720; CG8312.
 DR InterPro; IPR006020; PTB_PID.
 DR SMART; SM00462; PTB; 1_
 SQ SEQUENCE 569 AA; 63587 MW; 4463ECB6092EAE59 CRC64;

Query Match 90.2%; Score 37; DB 5; Length 569;
 Best Local Similarity 100.0%; Pred. No. 56;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNYRP 6
 Db 425 GNNYRP 430
 |||||

RESULT 5

Q9DE61 ID Q9DE61 PRELIMINARY; PRT; 742 AA.
 AC Q9DE61;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE P120 (Fragment).

OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 OC Xenopodinae; Xenopus.

OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ciesiolka M., Vanlandschoot A., Staes K., van Roy F.;
 RT "Armaddillo-related proteins in Xenopus laevis";
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF150746; AAG45945.1; -
 DR InterPro; IPR008938; ARM.
 DR InterPro; IPR000225; Armadillo.
 DR Pfam; PF00514; Armadillo_seg; 4.
 DR SMART; SM00185; ARM; 3.
 DR PROSITE; PS00176; ARM_REPEAT; 3.
 ET NON_TER 1
 SQ SEQUENCE 742 AA; 82927 MW; E51D3A3A5C895FE7 CRC64;

Query Match 90.2%; Score 37; DB 13; Length 742;
 Best Local Similarity 100.0%; Pred. No. 74;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNYRP 6
 Db 28 GNNYRP 33
 |||||

RESULT 6

Q9VHA6 ID Q9VHA6 PRELIMINARY; PRT; 866 AA.
 AC Q9VHA6;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE CG8312 protein.
 GN CG8312.

OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkeley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celinker S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA Amanatides P.G., Scherer S.E., Ashburner M., Henderson S.N.,
 RA George R.A., Lewis S.E., Richards S., Zhang Q., Chen L.X.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Blazef R.G., Champ M., Pfeiffer B.D.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mout S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*";
RL Science 287:2185-2195 (2000).
RN [2]
RN SEQUENCE FROM N.A.
RA Celnik S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Anantides P.G., Brandon R.C., Rogers Y.,
RA Banazon J., An H., Baldwin D., Benson J., Beeson K.Y., Busam D.A.,
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA Ferrera S., Frise E., Galle R.F., Gaig N.S., George R.A.,
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwam C., Jalali M., Kuse D., Li P., Mattei B., Moshrefi A.,
RA Mcintosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT "Sequencing of *Drosophila melanogaster* genome";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RN SEQUENCE FROM N.A.
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,
RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celnik S.E.,
RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA Krommiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT "Annotation of *Drosophila melanogaster* genome";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RN SEQUENCE FROM N.A.
RA Adams M.D., Celnik S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RN SEQUENCE FROM N.A.
RA FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB003683; AAF54413.2; --
DR FlyBase; FBgn0037720; CG8312.
DR InterPro; IPR006020; PTB_PID.
DR SMART; SM00462; PTB; 1.
SQ SEQUENCE 866 AA; 95296 MW; E1B0226E79A64A05 CRC64;
Query Match 90.2%; Score 37; DB 5; Length 866;
Best Local Similarity 100.0%; Pred. No. 87;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GNNYRP 6
DB 425 GNNYRP 430
RESULT 7
ID Q9DFB2 PRELIMINARY; PRT; 907 AA.
AC Q9DFB2;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Catenin arcfv-2ABC protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RN SEQUENCE FROM N.A.

RX MEDLINE=20459074; PubMed=10899158;
RA Paulson A.F., Mooney E., Fang X., Ji H., McCrea P.D.;
RT "Xarvcf: Xenopus Member of the p120 Catenin Subfamily Associating with
RT Cadherin Juxtamembrane Region";
RL J. Biol. Chem. 275:30124-30131 (2000).
DR EMBL; AF287051; AAG00555.1; --
DR InterPro; IPR008938; ARM.
DR InterPro; IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 4.
DR SMART; SM00185; ARM; 3.
DR PROSITE; PS00176; ARM_REPEAT; 3.
SQ SEQUENCE 907 AA; 101572 MW; A3B9EF7A0E5B9B86 CRC64;
Query Match 90.2%; Score 37; DB 13; Length 907;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GNNYRP 6
DB 155 GNNYRP 160
RESULT 8
ID Q9NS70 PRELIMINARY; PRT; 551 AA.
AC Q9NS70;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Transcription factor BMAL2.
GN BMAL2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=2042287; PubMed=10964693;
RA Ikeda M., Yu W., Hirai M., Ebisawa T., Honma S., Yoshimura K.,
RA Honma K., Nomura M.;
RT "cDNA cloning of a novel bHLH-PAS transcription factor superfamily
RT gene, BMAL2; its mRNA expression, subcellular distribution, and
RT chromosomal localization";
RL Biochem. Biophys. Res. Commun. 275:493-502 (2000).
CC -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS.
DR EMBL; AB039921; BAB01485.1; --
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0004871; F:signal transducer activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR001092; HLH_basid.
DR InterPro; IPR001067; Nuc_translocat.
DR InterPro; IPR001610; PAC_
DR InterPro; IPR000014; PAS_domain.
DR Pfam; PF00010; HLH; 1.
DR Pfam; PF00989; PAS; 2.
DR PRINTS; PR00785; NCTRNLOCATR.
DR SMART; SM00353; HLH; 1.
DR SMART; SM00086; PAC; 1.
DR SMART; SM00091; PAS; 2.
DR TIGRfam; TIGR00229; sensory_box; 2.
DR PROSITE; PS00038; HLH_1; 1.
DR PROSITE; PS00888; HLH_2; 1.
DR PROSITE; PS00112; PAS; 2.
SQ SEQUENCE 551 AA; 61786 MW; FF93B29B660E218D CRC64;
Query Match 87.8%; Score 36; DB 4; Length 551;
Best Local Similarity 85.7%; Pred. No. 85;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
QY      1 GNNYRPS 7
Db      84 GSNYRPS 90

RESULT 9
Q96J63 PRELIMINARY; PRT; 551 AA.
AC Q96J63;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Similar to transcription factor BMAL2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Strausberg R.;

Query Match 87.8%; Score 36; DB 4; Length 551;
Best Local Similarity 85.7%; Pred. No. 85;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

CC -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
TRANSCRIPTION FACTORS.
DR EMBL; BC000172; AAH00172.2; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0004871; F:signal transducer activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR001092; HLH_basic.
DR InterPro; IPR001087; Nuc_translocat.
DR InterPro; IPR001610; PAC.
DR InterPro; IPR000014; PAS_domain.
DR Pfam; PF00010; HLH_1.
DR Pfam; PF00989; PAS; 2.
DR PRINTS; PR00785; NCTENSLCATR.
DR SMART; SM00353; HLH; 1.
DR SMART; SM00086; PAC; 1.
DR SMART; SM00091; PAS; 2.
DR TIGRFAMs; TIGR00229; sensory_box; 2.
DR PROSITE; PS00038; HLH_1; 1.
DR PROSITE; PS50888; HLH_2; 1.
DR PROSITE; PS50112; PAS; 2.
SQ SEQUENCE 551 AA; 61787 MW; F845E3F54FPA39B0 CRC64;

QY      1 GNNYRPS 7
Db      84 GSNYRPS 90

RESULT 10
Q9NYQ4 PRELIMINARY; PRT; 572 AA.
AC Q9NYQ4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE BHLH-PAS transcription factor MOP9.
GN MOP9.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=20325577; PubMed=10864977;
RA Hogenesch J.B., Gu Y.Z., Moran S.M., Shimomura K., Radcliffe L.A.,
RA Takahashi J.S., Bradfield C.A.;

Query Match 87.8%; Score 36; DB 4; Length 551;
Best Local Similarity 85.7%; Pred. No. 85;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

CC -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
TRANSCRIPTION FACTORS.
DR EMBL; AF231338; AAF71306.1; -.
DR HSSP; P36956; 1AM9.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0004871; F:signal transducer activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR001092; HLH_basic.
DR InterPro; IPR001087; Nuc_translocat.
DR InterPro; IPR001610; PAC.
DR InterPro; IPR000014; PAS_domain.
DR Pfam; PF00010; HLH; 1.
DR Pfam; PF00989; PAS; 2.

QY      1 GNNYRPS 7
Db      84 GSNYRPS 90

RESULT 11
Q9NYQ5 PRELIMINARY; PRT; 585 AA.
AC Q9NYQ5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE BHLH-PAS transcription factor MOP9.
GN MOP9.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=20325577; PubMed=10864977;
RA Hogenesch J.B., Gu Y.Z., Moran S.M., Shimomura K., Radcliffe L.A.,
RA Takahashi J.S., Bradfield C.A.;

Query Match 87.8%; Score 36; DB 4; Length 572;
Best Local Similarity 85.7%; Pred. No. 88;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 GNNYRPS 7
Db      105 GSNYRPS 111
```

```
RT "The basic helix-loop-helix-PAS protein MOP9 is a brain-specific
RT heterodimeric partner of circadian and hypoxia factors.";
RL J. Neurosci. 20:RC83-RC83(2000).
CC -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS.
DR EMBL; AF231339; AAF71307.1; -.
DR HSSP; P36956; 1AM9.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0004871; F:signal transducer activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR001092; HLH_basic.
DR InterPro; IPR001067; Nuc_translocat.
DR InterPro; IPR001610; PAC.
DR InterPro; IPR000014; PAS_domain.
DR Pfam; PF00010; HLH; 1.
DR Pfam; PF00989; PAS; 2.
DR PRINTS; PR00785; NCTENSLCATR.
DR SMART; SM00353; HLH; 1.
DR SMART; SM00086; PAC; 1.
DR SMART; SM00091; PAS; 2.
DR TIGRFAMs; TIGR00229; sensory_box; 2.
DR PROSITE; PS00038; HLH_1; 1.
DR PROSITE; PS50888; HLH_2; 1.
DR PROSITE; PS50112; PAS; 2.
SQ SEQUENCE 572 AA; 64162 MW; E4056207B02D1FD1 CRC64;

QY      1 GNNYRPS 7
Db      105 GSNYRPS 111

RESULT 11
Q9NYQ5 PRELIMINARY; PRT; 585 AA.
AC Q9NYQ5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE BHLH-PAS transcription factor MOP9.
GN MOP9.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=20325577; PubMed=10864977;
RA Hogenesch J.B., Gu Y.Z., Moran S.M., Shimomura K., Radcliffe L.A.,
RA Takahashi J.S., Bradfield C.A.;

Query Match 87.8%; Score 36; DB 4; Length 572;
Best Local Similarity 85.7%; Pred. No. 88;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 GNNYRPS 7
Db      105 GSNYRPS 111
```

```
DR PRINTS; PRO0785; NCTRNSLOCATR.
DR SMART; SM00353; HLH; 1.
DR SMART; SM00086; PAC; 1.
DR SMART; SM00091; PAS; 2.
DR TIGRFAMs; TIGR00229; sensory_box; 2.
DR PROSITE; PS00038; HLH_1; 1.
DR PROSITE; PS0888; HLH_2; 1.
DR PROSITE; PS0112; PAS; 2.
SQ SEQUENCE 585 AA; 65556 MW; 7F744D78BB8ADF15 CRC64;

Query Match      87.8%; Score 36; DB 4; Length 585;
Best Local Similarity 85.7%; Pred. No. 91;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 GNNYRPS 7
Db      118 GSNYRPS 124

RESULT 12
Q8WYA1 PRELIMINARY; PRT; 588 AA.
AC Q8WYA1;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Brain-muscle-ARNT-like transcription factor 2d.
GN BMAL2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Okano T., Sasaki M., Yamamoto K., Hirota T., Takanaka Y.,
RA Fukada Y.;
RT "Identification of a novel ARNT-like transcription factor, hBMAL2.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
CC -|- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS.
DR EMBL; AF246963; AAL50342.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0004871; P:signal transducer activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR001067; HLH_basic.
DR Brain-muscle-ARNT-like transcription factor 2d.
GN BMAL2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Okano T., Sasaki M., Yamamoto K., Hirota T., Takanaka Y.,
RA Fukada Y.;
RT "Identification of a novel ARNT-like transcription factor, hBMAL2.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
CC -|- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS.
DR EMBL; AF246963; AAL50342.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0004871; P:signal transducer activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR001067; HLH_basic.
DR InterPro; IPR001092; Nuc_translocat.
DR InterPro; IPR001610; PAC_domain.
DR Pfam; PF00010; HLH; 1.
DR Pfam; PF00989; PAS; 2.
DR PRINTS; PRO0785; NCTRNSLOCATR.
DR SMART; SM00353; HLH; 1.
DR SMART; SM00086; PAC; 1.
DR SMART; SM00091; PAS; 2.
DR TIGRFAMs; TIGR00229; sensory_box; 2.
DR PROSITE; PS00038; HLH_1; 1.
DR PROSITE; PS0888; HLH_2; 1.
DR PROSITE; PS0112; PAS; 2.
SQ SEQUENCE 599 AA; 66493 MW; A5B1B84DBC4B5B40 CRC64;

Query Match      87.8%; Score 36; DB 4; Length 599;
Best Local Similarity 85.7%; Pred. No. 93;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 GNNYRPS 7
Db      132 GSNYRPS 138

RESULT 14.
Q8WYA2 PRELIMINARY; PRT; 599 AA.
AC Q8WYA2;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Brain-muscle-ARNT-like transcription factor 2c.
GN BMAL2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Okano T., Sasaki M., Yamamoto K., Hirota T., Takanaka Y.,
RA Fukada Y.;
RT "Identification of a novel ARNT-like transcription factor, hBMAL2.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
CC -|- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS.
DR EMBL; AF246962; AAL50341.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0004871; P:signal transducer activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR001067; HLH_basic.
DR InterPro; IPR001092; Nuc_translocat.
DR InterPro; IPR001610; PAC_domain.
DR InterPro; IPR000014; PAS_domain.
DR Pfam; PF00010; HLH; 1.
DR Pfam; PF00989; PAS; 2.
DR PRINTS; PRO0785; NCTRNSLOCATR.
DR SMART; SM00353; HLH; 1.
DR SMART; SM00086; PAC; 1.
DR SMART; SM00091; PAS; 2.
DR TIGRFAMs; TIGR00229; sensory_box; 2.
DR PROSITE; PS00038; HLH_1; 1.
DR PROSITE; PS0888; HLH_2; 1.
DR PROSITE; PS0112; PAS; 2.
SQ SEQUENCE 599 AA; 66493 MW; A5B1B84DBC4B5B40 CRC64;

Query Match      87.8%; Score 36; DB 4; Length 599;
Best Local Similarity 85.7%; Pred. No. 93;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 GNNYRPS 7
Db      132 GSNYRPS 138

RESULT 14.
Q8WYA2 PRELIMINARY; PRT; 602 AA.
AC Q8WYA2;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Cycle-like factor CLIF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Medline=20538426; PubMed=11018023;
RA Memura K., de La Monte S.M., Chin M.T., Layne M.D., Hsieh C.M.,
RA Vet S.F., Perrella M.A., Lee M.E.;
RT "CLIF, a Novel Cycle-like Factor, Regulates the Circadian Oscillation
RT of Plasmidogen Activator Inhibitor-1 Gene Expression.";
RL J. Biol. Chem. 275:36847-36851(2000).
CC -|- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS.
DR EMBL; AF256215; AAG34652.1; -.

```



```

OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Matthews L.;
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL683828; CAD62265.1; -.
DR PIR; PT0546; PT0698.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005179; F:hormone activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0007275; P:development; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR002012; GNRH.
DR PROSITE; PS00473; GNRH; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
SQ SEQUENCE 1404 AA; 153123 MW; D24B6545BFA01372 CRC64;

Query Match      87.8%; Score 36; DB 11; Length 1404;
Best Local Similarity 85.7%; Pred. No. 2.3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 GNNYRPS 7
DB      164 GSNYRPS 170

RESULT 18
Q8J0D5
ID Q8J0D5 PRELIMINARY; PRT; 274 AA.
AC Q8J0D5;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Orf136.
GN Orf136.
OS Heliothis zea virus 1.
OC Viruses; dsDNA viruses, no RNA stage; unclassified dsDNA viruses.
OX NCBI_TaxID=29250;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22174892; PubMed=12186886;
RA Cheng C.H., Liu S.M., Chow T.Y., Hsiao Y.Y., Wang D.P., Huang J.J.,
  Chen H.H.;
RT "Analysis of the Complete Genome Sequence of the H2-1 Virus Suggests
  that It is Related to Members of the Baculoviridae.";
RL J. Virol. 76:9024-9034(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Cheng C.H., Liu H.M., Hsiao Y.Y., Chow T.Y., Chen H.H.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Cheng C.H., Hsiao Y.Y., Liu S.M., Chow T.Y., Chen H.H.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF451898; RAN04428.1; -.
DR InterPro; IPR001202; WW_Rsp5_WNP.
DR PROSITE; PS01159; WW_DOMAIN_1; 1.
SQ SEQUENCE 274 AA; 30555 MW; 978CE42F9FFCA93F CRC64;

Query Match      85.4%; Score 35; DB 12; Length 274;
Best Local Similarity 85.7%; Pred. No. 65;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GNNYRPS 7
DB      175 GNNYRPS 181

RESULT 19
Q8J0D5
ID Q8J0D5 PRELIMINARY; PRT; 632 AA.
AC Q8J0D5;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Metalloprotease, MEPI precursor.
GN MEPI.
OS Microsporium canis.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Onygenales; Arthrodermataceae; mitosporic Arthrodermataceae;
OC Microsporium.
OX NCBI_TaxID=82078;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22215703; PubMed=12228297;
RA Brouta F., Descamps F., Monod M., Vermout S., Losson B., Mignon B.;
RT "Secreted metalloprotease gene family of Microsporium canis.";
RL Infect. Immun. 70:5676-5683(2002).
DR EMBL; AJ490184; CAD35289.1; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001842; Peptidase_M36.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR Pfam; PF02128; Peptidase_M36; 1.
DR PRINTS; PR00999; FUNGALYSIN.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Signal; Protease.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 244 632 METALLOPROTEASE, MEPI.
SQ SEQUENCE 632 AA; 69770 MW; D2FF2E8E8FABEBAF CRC64;

Query Match      85.4%; Score 35; DB 3; Length 632;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 NNYRPS 7
DB      307 NNYRPS 312

RESULT 20
Q8NIB6
ID Q8NIB6 PRELIMINARY; PRT; 635 AA.
AC Q8NIB6;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative secreted metalloprotease 1.
GN MEPI.
OS Trichophyton rubrum.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Onygenales; Arthrodermataceae; mitosporic Arthrodermataceae;
OC Trichophyton.
OX NCBI_TaxID=5551;
RN [1]
RP SEQUENCE FROM N.A.
RA Capocchia S., Lechenne B., Zaugg C., Monod M.;
RT "Trichophyton rubrum encoding metalloprotease.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF407185; RAN03636.1; -.
DR EMBL; AF407186; RAN03637.1; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0008270; F:metalloendopeptidase activity; IEA.
DR GO; GO:0004222; F:zinc ion binding; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001842; Peptidase_M36.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR Pfam; PF02128; Peptidase_M36; 1.
DR PRINTS; PR00999; FUNGALYSIN.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Metalloprotease; Protease.
SQ SEQUENCE 635 AA; 70126 MW; EA1123969FC0DB19 CRC64;

```

```

Query Match      85.4%; Score 35; DB 3; Length 635;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 NNYRPS 7
DB      310 NNYRPS 315

RESULT 21
Q8PUB4      PRELIMINARY; PRT; 982 AA.
AC Q8PUB4;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Chaperone protein.
GN HTPG OR M2421.
OS Methanosarcina mazei (Methanosarcina frisia).
OC Archaea; Euryarchaeota; Euryarchaeota orders incertae sedis;
OC Methanosarcinales; Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2209;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Goel / Go1 / ATCC BAA-199 / DSM 3647 / OCM 88;
RX MEDLINE=22120827; PubMed=12125824;
RA Deppenmeier U., Johann A., Hartsch T., Merkl R., Schmitz R.A.,
RA Martinez-Arias R., Henne A., Wietzer A., Baeumer S., Jacobi C.,
RA Brueggemann H., Lienard T., Christmann A., Boemcke M., Steckel S.,
RA Bhattacharyya A., Lykidis A., Overbeek R., Klenk H.-P., Gunsalus R.P.,
RA Fritz H.-J., Gottschalk G.;
RT "The genome of Methanosarcina mazei: evidence for lateral gene
RT transfer between Bacteria and Archaea.";
RL J. Mol. Microbiol. Biotechnol. 4:453-461(2002).
DR EMBL; A013485; AAM32117.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003754; F:chaperone activity; IEA.
DR InterPro; IPR003594; ATPbind_ATPase.
DR InterPro; IPR001404; Hsp90.
DR Pfam; PF02518; HATPase C; 1.
DR PRINTS; PR00775; HEATSHOCK90.
KW Complete proteome.
SQ SEQUENCE .982 AA; 114565 MW; 24B3340A494998639 CRC64;

Query Match      85.4%; Score 35; DB 17; Length 982;
Best Local Similarity 85.7%; Pred. No. 2.5e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GNNYRPS 7
DB      395 GNNYRPS 401

RESULT 22
Q87TJ7      PRELIMINARY; PRT; 1046 AA.
AC Q87TJ7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Sensory box/GGDEF family protein.
GN VP0071.
OS Vibrio parahaemolyticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=670;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RMD 2210633 / Serotype O3:K6;
RX MEDLINE=22508454; PubMed=12620739;
RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
RA Iijima Y., Najima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,

```

```

RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
RT distinct from that of V. cholerae.";
RL Lancet 361:743-749(2003).
DR EMBL; AP005073; BAC58335.1; -.
DR GO; GO:0000155; F:two-component sensor molecule activity; IEA.
DR GO; GO:0000160; P:two-component signal transduction system (p. . .; IEA.
DR InterPro; IPR001633; EAL.
DR InterPro; IPR000160; GGDEF.
DR InterPro; IPR000160; PAC.
DR InterPro; IPR000700; PAS-assoc C.
DR InterPro; IPR000114; PAS domain.
DR InterPro; IPR000437; Prok_lipoprot_S.
DR Pfam; PF00563; EAL; 1.
DR Pfam; PF00990; GGDEF; 1.
DR Pfam; PF00785; PAC; 2.
DR PROSITE; PS00883; EAL; 1.
DR PROSITE; PS00887; GGDEF; 1.
DR PROSITE; PS01113; PAC; 1.
DR PROSITE; PS01112; PAS; 1.
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW Complete proteome.
SQ SEQUENCE 1046 AA; 120524 MW; 918AB439BDA9DB38 CRC64;

Query Match      85.4%; Score 35; DB 16; Length 1046;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 NNYRPS 7
DB      123 NNYRPS 128

RESULT 23
Q8EX31      PRELIMINARY; PRT; 1346 AA.
AC Q8EX31;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN MYPE190.
OS Mycoplasma penetrans.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=28227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HF-2;
RX MEDLINE=22354719; PubMed=12466555;
RA Sasaki Y., Ishikawa J., Yamashita A., Oshima K., Kenri T., Furuya K.,
RA Yoshino C., Horino A., Shiba T., Sasaki T., Hattori M.;
RT "The complete genomic sequence of Mycoplasma penetrans, an
RT intracellular bacterial pathogen in humans.";
RL Nucleic Acids Res. 30:5293-5300(2002).
DR EMBL; AP004170; BAC43809.1; -.
DR Hypothetical protein; Complete proteome.
SQ SEQUENCE 1346 AA; 146839 MW; A5B2E8FA4026761F CRC64;

Query Match      85.4%; Score 35; DB 16; Length 1346;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 NNYRPS 7
DB      377 NNYRPS 382

RESULT 24
O96240      PRELIMINARY; PRT; 1377 AA.
AC O96240;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-2003 (TrEMBLrel. 23, Last sequence update)

```

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN PFB0735C.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3D7;
RX MEDLINE=99021743; PubMed=9804551;
RA Gardner M.J., Tetrelin H., Carucci D.J., Cummings L.M., Aravind L.,
RA Koonin E.V., Shallom S., Mason T., Yu K., Fujii C., Pederson J.,
RA Shen K., Jing J., Aston C., Lai Z., Schwartz D.C., Pertea M.,
RA Salzberg S., Zhou L., Sutton G.G., Clayton R., White O., Smith H.O.,
RA Fraser C.M., Adams M.D., Venter J.C., Hoffman S.L.;
RT "Chromosome 2 sequence of the human malaria parasite Plasmodium
falciparum";
RL Science 282:1126-1132(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=3D7;
RX MEDLINE=22255705; PubMed=12368864;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.W., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shallom S.J., Suh B., Peterson J., Angiuoli S.,
RA Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McPadden G.I., Cummings L.M., Subramanian G.M., Murgall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrell B.;
RT "Genome sequence of the human malaria parasite Plasmodium
falciparum";
RL Nature 419:498-511(2002).
DR EMBL; AE001414; AAC71936.2; --
DR FIR; G71607; G71607.
KW Hypothetical protein.
SQ SEQUENCE 1377 AA; 167626 MW; 99E2C9D0D9E0B7A8 CRC64;

Query Match 85.4%; Score 35; DB 5; Length 1377;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 NNYRPS 7
DB 819 NNYRPS 824
|||||
|||

RESULT 25
O7X3P2
ID O7X3P2 PRELIMINARY; PRT; 378 AA.
AC O7X3P2;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Integrase protein.
GN SIP.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=V329;
RX MEDLINE=21189316; PubMed=11292810;
RA Cucarella C., Solano C., Valle J., Amorena B., Lasa I., Penades J.R.;
RT "Bap, a Staphylococcus aureus surface protein involved in biofilm
formation";
RT J. Bacteriol. 183:2888-2896(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=V329;
RA Ubeda C., Tormo M.A., Cucarella C., Trotonda P., Foster T.J., Lasa I.,
RA Penades J.R.;

RT "Sip, an integrase protein with excision, circularization and
RT integration activities, defines a new family of mobile Staphylococcus
RT aureus pathogenicity islands.";
RL Mol. Microbiol. 0:0-0(2003).
DR EMBL; AY220730; AAP5251.1; --
SQ SEQUENCE 378 AA; 44587 MW; CAFD7513DF8C960E CRC64;

Query Match 82.9%; Score 34; DB 2; Length 378;
Best Local Similarity 83.3%; Pred. No. 1.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GNNYRP 6
DB 233 GNNYRP 238
|||||
|||

Search completed: September 24, 2004, 02:18:33
Job time : 42.5472 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2005, 17:57:45 ; Search time 45.2812 Seconds
(without alignments)

79.162 Million cell updates/sec

Title: US-10-088-639a-2_COPY_49_55
Perfect score: 41
Sequence: 1 GNNYRPS 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37	90.2	146	2 Q84KK9	Q84KK9 phlebotomium
2	37	90.2	327	2 Q9W316	Q9W316 drosophila
3	37	90.2	447	2 Q86BA6	Q86BA6 drosophila
4	37	90.2	569	2 Q8WRF2	Q8WRF2 drosophila
5	37	90.2	742	2 Q9DE61	Q9DE61 xenopus lae
6	37	90.2	866	2 Q9VHA6	Q9VHA6 drosophila
7	37	90.2	868	2 Q6INZ6	Q6INZ6 xenopus lae
8	37	90.2	907	2 Q9DPB2	Q9DPB2 xenopus lae
9	37	90.2	951	1 XR22 MOUSE	Q9DBR1 mus musculus
10	37	90.2	1369	2 Q756Z3	Q756Z3 ashbya goss
11	36	87.8	387	1 SSUD_BRAJA	Q89ER2 brachyrrhizob
12	36	87.8	551	2 Q9NS70	Q9NS70 homo sapien
13	36	87.8	572	2 Q9NYQ4	Q9NYQ4 homo sapien
14	36	87.8	578	2 Q96J63	Q96J63 homo sapien
15	36	87.8	585	2 Q9NYO5	Q9NYO5 homo sapien
16	36	87.8	588	2 Q8WYA1	Q8WYA1 homo sapien
17	36	87.8	599	2 Q8WYA2	Q8WYA2 homo sapien
18	36	87.8	602	2 Q9H2M4	Q9H2M4 homo sapien
19	36	87.8	611	2 Q72HQ6	Q72HQ6 thermus the
20	36	87.8	622	2 Q8WYA3	Q8WYA3 homo sapien
21	36	87.8	636	2 Q8WYA4	Q8WYA4 homo sapien
22	36	87.8	1100	2 Q6ZPF9	Q6ZPF9 mus musculus
23	36	87.8	1404	2 Q80VW7	Q80VW7 mus musculus
24	35	85.4	222	1 T2M1_METJA	Q58391 methanococ
25	35	85.4	274	2 Q8JKH7	Q8JKH7 heliothis z
26	35	85.4	632	2 Q8JOD5	Q8JOD5 microsporim
27	35	85.4	634	1 ELM1_ASPFU	P46074 aspergillus
28	35	85.4	634	1 ELM2_ASPFU	P46075 aspergillus
29	35	85.4	635	2 Q8NIB6	Q8NIB6 trichophyto
30	35	85.4	635	2 Q6WIH5	Q6WIH5 arthroderma
31	35	85.4	798	2 Q6GIJ3	Q6GIJ3 bartonella

32 35 85.4 982 2 Q8PUB4
33 35 85.4 1046 2 Q87TJ7
34 35 85.4 1346 2 Q8EX31
35 35 85.4 1377 2 Q96240
36 34 82.9 378 2 Q7WS19
37 34 82.9 378 2 Q7X3P2
38 34 82.9 461 2 Q6ST68
39 34 82.9 461 2 Q6ST73
40 34 82.9 462 2 Q6ST65
41 34 82.9 465 2 Q6ST72
42 34 82.9 872 2 Q7QRG1
43 34 82.9 919 1 YB03_YEAST
44 33 80.5 292 2 Q7QEE7
45 33 80.5 311 2 Q39500

Q8pub4 methanosarc
Q87tj7 vibrio para
Q8ex31 mycoplasma
Q96240 plasmodium
Q7ws19 staphylococ
Q7x3p2 staphylococ
Q6st68 human rhino
Q6st73 human rhino
Q6st65 human rhino
Q6st72 human rhino
Q7qrg1 giardia lam
P38073 saccharomyc
Q7qee7 anopheles g
Q39500 cylindrothe

ALIGNMENTS

RESULT 1
Q84KK9 PRELIMINARY; PRT; 146 AA.
AC Q84KK9
DT 01-JUN-2003 (Tremblrel. 24, Created)
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Jacalin-related lectin.
GN Name=PALB;
OS Phlebotomium aureum.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Moniliformopses; Filicophyta; Filicopsida; Filicales; Polypodiaceae;
OC Phlebotomium.
OX NCBI_TaxID=218620;
RN [1]_SEQUENCE FROM N.A.
RP TISSUE=Rhizome;
RX MEDLINE=22538482; PubMed=12538584; DOI=10.1074/jbc.M211840200;
RA Tateno H., Harry W.C., Petryniak J., Goldstein I.J.;
RT "Purification, characterization, molecular cloning, and expression of
RT novel members of jacalin-related lectins from rhizomes of the true
RT fern Phlebotomium aureum (L) J. Smith (Polypodiaceae).";
RL J. Biol. Chem. 278:10891-10899(2003).
DR EMBL; AB099933; BAC55269.1; .
DR GO; GO:0005529; F:sugar binding; IEA.
DR InterPro; IPR001229; Jacalin_lectin.
KW Lectin.
SQ SEQUENCE 146 AA; 15026 MW; 55353B28E1ECEA6A CRC64;

Query Match 90.2%; Score 37; DB 2; Length 146;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNYRP 6
DB 117 GNNYRP 122

RESULT 2
Q9W316 PRELIMINARY; PRT; 327 AA.
ID Q9W316
AC Q9W316; Q961F5;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 25-OCT-2004 (Tremblrel. 28, Last annotation update)
DE CG32698-PA (GH26663P).
GN Name=CG9683; ORFNames=CG32698;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]_SEQUENCE FROM N.A.
RP SEQUENCE FROM N.A.

RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA April J.F., Agbayani A., An H.J., Andrews-Pfannkuch C., Baldwin D.,
RA Ballaw R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleeb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [12]
RN SEQUENCE FROM N.A.
RX MEDLINE=2426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,
RA Pacleeb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskaas R., Tabor P.E., Wan K.H., Stapleton M., Sutton G.G., Venter C.,
RA Weinstein G., Scher S.E., Myers E.W., Gibbs R.A., Rubin G.M.,
RT "Finishing a whole-genome shotgun: Release 3 of the *Drosophila*
RT melanogaster euchromatic genome sequence.";
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [13]
RN SEQUENCE FROM N.A.
RX MEDLINE=2426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskaas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the *Drosophila melanogaster* euchromatin:
RT a genomics perspective.";
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
RN [14]
RN SEQUENCE FROM N.A.
RX MEDLINE=2426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.O.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a
RT systematic review.";
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).

RN [5]
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nuccio J., Pacleeb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF003448; AAF46521.2; -;
DR EMBL; AY051620; AAK93044.1; -;
DR HSSP; P00918; 1BV3.
DR FlyBase; FBgn0052698; CG32698.
DR GO; GO:0004089; F:carbonate dehydratase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006730; P:one-carbon compound metabolism; IEA.
DR InterPro; IPR001148; Euk_Coanhd.
DR Pfam; PF00194; Carb_anhydase; 1.
SQ SEQUENCE 327 AA; 37182 MW; 9574C0BB8BA9FD01 CRC64;
Query Match 90.2%; Score 37; DB 2; Length 327;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GNNYRP 6
Db 279 GNNYRP 284
|||||
RESULT 3
Q86BA6 PRELIMINARY; PRT; 447 AA.
ID Q86BA6
AC Q86BA6;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE CG8312-PB.
GN ORFNames=CG8312;
OS *Drosophila melanogaster* (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; *Drosophila*.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=10731132; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA April J.F., Agbayani A., An H.J., Andrews-Pfannkuch C., Baldwin D.,
RA Ballaw R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,

QY 1 GNNYRP 6
28 GNNYRP 33
Db

```

RESULT 6
Q9VHA6 PRELIMINARY; PRT; 866 AA.
AC Q9VHA6;
DT 01-MAR-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE CG8312-PA.
OS ORFNames=CG8312;
GN Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA April J.F., Abayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burris K.C., Busam D.A., Butler H., Cadenot E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.B., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirekas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wasarman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., WoodageT, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richmond S., Sodergren E.J.,
RA Svirekas R., Taber P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstock G., Scher S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila
RT melanogaster euchromatic genome sequence."
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirekas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,

```

```

RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
RT a Genomics perspective.";
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.O.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review.";
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RN [5]
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003683; RAP54413.2; -.
DR FlyBase; FBgn0037720; CG8312.
DR InterPro; IPR011036; PH-related.
SQ SEQUENCE 866 AA; 95236 MW; E1B0226E79A64A05 CRC64;
Query Match 90.2%; Score 37; DB 2; Length 866;
Best Local Similarity 100.0%; Pred. No. 94; Mismatches 0; Indels 0; Gaps 0;
Matches 6; Conservative 0;
QY 1 GNNYRP 6
DB 425 GNNYRP 430
|||||
RESULT 7
Q6INZ6 PRELIMINARY; PRT; 868 AA.
AC Q6INZ6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Roschlyuk S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mulhany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huiyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whitling M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;

```

RT "Generation and initial analysis of more than 15,000 full-length human
 RL and mouse cDNA sequences.",
 RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002). [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Splicein;
 RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
 RA Richardson P.;
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 RT initiative",
 RL Dev. Dyn. 225:384-391(2002). [3]

RN SEQUENCE FROM N.A.

RP TISSUE=Splicein;
 RC TISSUE=Splicein;
 RA Klein S., Strausberg R.;
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC071224; AAH7124.1; -;

DR InterPro; IPR008938; ARM.

DR InterPro; IPR000225; Armadillo.

DR Pfam; PF00514; Arm; 4.

DR SMART; SM00185; ARM; 6.

DR PROSITE; PS0176; ARM_REPEAT; 3.

KW Hypothetical protein.

SQ SEQUENCE 868 AA; 96907 MW; 433B1B5FB6CCF08F CRC64;

Query Match 90.2%; Score 37; DB 2; Length 868;

Best Local Similarity 100.0%; Pred. No. 94;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GNNYRP 6

Db 154 GNNYRP 159

RESULT 8

Q9DFB2

ID Q9DFB2 PRELIMINARY; PRT; 907 AA.

AC Q9DFB2

DT 01-WAR-2001 (TREMELrel. 16, Created)

DT 01-WAR-2001 (TREMELrel. 16, Last sequence update)

DT 01-WAR-2001 (TREMELrel. 26, Last annotation update)

DE Catenin arvcf-2ABC protein.

OS Xenopus laevis (African clawed frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;

OC Xenopodinae; Xenopus.

OX NCBI_TaxID=8355;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=20459074; PubMed=10899158; DOI=10.1074/jbc.M003048200;

RA Paulson A.F., Mooney E., Fang X., Ji H., McCrea P.D.;

RT "Xarvcf: Xenopus Member of the p120 Catenin Subfamily Associating with

RL Caderin Juxtamembrane Region.",

RL J. Biol. Chem. 275:30124-30131(2000). [1]

DR EMBL; AF287051; AAG00555.1; -;

DR InterPro; IPR008938; ARM.

DR InterPro; IPR000225; Armadillo.

DR Pfam; PF00514; Arm; 4.

DR SMART; SM00185; ARM; 5.

DR PROSITE; PS0176; ARM_REPEAT; 3.

SQ SEQUENCE 907 AA; 101572 MW; A3B9EF7A0E5B9B86 CRC64;

Query Match

Best Local Similarity 90.2%; Score 37; DB 2; Length 907;

Matches 6; Conservative 0; Mismatches 0; Indels 0;

Qy 1 GNNYRP 6

Db 155 GNNYRP 160

RESULT 9

XRN2_MOUSE
 ID XRN2_MOUSE STANDARD; PRT; 951 AA.
 AC Q9DBR1; O61489; O99KS7;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE 5'-3' exoribonuclease 2 (EC 3.1.11.-) (Dhml protein).
 GN Name=Xrn2; Synonyms=Dhml;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC TISSUE=Testis;
 RX MEDLINE=95192042; PubMed=7885830;
 RA Shobuie T., Sugano S., Yanashita T., Ikeda H.;
 RT "Characterization of cDNA encoding mouse homolog of fission yeast
 RL dhpl+ gene: structural and functional conservation.",
 RN Nucleic Acids Res. 23:357-361(1995). [2]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC STRAIN=CS5BL/6J; TISSUE=Embryo, Eye, Forelimb, and Lung;
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nikaide I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojohori T.,
 RA Baladrelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schriml L.M., Knapin A., Matsuda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Bradt D., Bruscia V., Chothia C., Corbani L.E., Cousins S.,
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sandelin A., Schneider C., Sempile C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Tesdale R.D., Tomita M.,
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.",
 RL Nature 420:563-573(2002). [3]
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC STRAIN=CS5BL/6, and Czech II; TISSUE=Brain, and Mammary Gland;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullah S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,

or send an email to license@isb-sib.ch).

CC EMBL; AP005960; BAC52275.1; --
DR HSSP; P80645; 1M41.
DR HAMAP; MF 01229; --; 1.
DR InterPro; IPR002103; Bac_luciferase.
DR InterPro; IPR011251; Luciferase_like.
DR Pfam; PF00296; Bac_luciferase; 1.
DR Complete proteome; FMN; Monooxygenase; Oxidoreductase.
KW SEQUENCE 387 AA; 42411 MW; C63361B094AD8431 CRC64;

Query Match 87.8%; Score 36; DB 1; Length 387;

Best Local Similarity 85.7%; Pred. No. 66; Indels 0; Gaps 0;
Matches 6; Conservative 1; Mismatches 0;

QY 1 GNNYRPS 7

||:||||

Db 375 GNDYRPS 381

RESULT 12

ID Q9NS70 PRELIMINARY; PRT; 551 AA.
AC Q9NS70; MEDLINE=20422287; PubMed=10964693; DOI=10.1006/bbr.2000.3248;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE Transcription factor BMAL2.
GN Name=Bmal2;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1] SEQUENCE FROM N.A.
RP TISSUE=Brain;
RC MEDLINE=20422287; PubMed=10964693; DOI=10.1006/bbr.2000.3248;
RA Ikeda M., Yu W., Hirai M., Ebisawa T., Honma S., Yoshimura K.,
RA Honma K., Nomura M.;
RT "cDNA cloning of a novel bHLH-PAS transcription factor superfamily
RT gene, BMAL2; its mRNA expression, subcellular distribution, and
RL Biochem. Biophys. Res. Commun. 275:493-502 (2000).
CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
DR EMBL; AB039921; BAB01485.1; --
DR HSSP; Q99814; 1P97.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0004871; F:signal transducer activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR001092; HLH_basic.
DR InterPro; IPR001067; Nuc_translocat.
DR InterPro; IPR000014; PAS.
DR Pfam; PF00010; HLH; 1.
DR Pfam; PF00989; PAS; 2.
DR PRINTS; PR00785; NCTRNLOCATR.
DR SMART; SM00353; HLH; 1.
DR SMART; SM00086; PAC; 1.
DR SMART; SM00091; PAS; 2.
DR PROSITE; PS50888; HLH; 1.
DR PROSITE; PS50112; PAS; 2.
DR PROSITE; PS50112; PAS; 2.
SQ SEQUENCE 551 AA; 61786 MW; PF93B29B660E218D CRC64;

Query Match 87.8%; Score 36; DB 2; Length 551;

Best Local Similarity 85.7%; Pred. No. 94; Indels 0; Gaps 0;
Matches 6; Conservative 1; Mismatches 0;

QY 1 GNNYRPS 7

||:||||

Db 84 GSNYRPS 90

RESULT 13

Q9NYQ4 PRELIMINARY; PRT; 572 AA.
ID Q9NYQ4;
AC Q9NYQ4; MEDLINE=20325577; PubMed=10864977;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE BHLH-PAS transcription factor MOP9.
GN Name=MOP9;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1] SEQUENCE FROM N.A.
RP MEDLINE=20325577; PubMed=10864977;
RX Hogenesch J.B., Gu Y.Z., Moran S.M., Shimomura K., Radcliffe L.A.,
RA Takahashi J.S., Bradfield C.A.;
RT "The basic helix-loop-helix-PAS protein MOP9 is a brain-specific
RT heterodimeric partner of circadian and hypoxia factors.";
RL J. Neurosci. 20:RC83-RC83 (2000).
CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
DR EMBL; AF231339; AAF71307.1; --
DR HSSP; Q99814; 1P97.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0004871; F:signal transducer activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR001092; HLH_basic.
DR InterPro; IPR001067; Nuc_translocat.
DR InterPro; IPR001610; PAC.
DR InterPro; IPR000014; PAS.
DR Pfam; PF00010; HLH; 1.
DR Pfam; PF00989; PAS; 2.
DR PRINTS; PR00785; NCTRNLOCATR.
DR SMART; SM00353; HLH; 1.
DR SMART; SM00086; PAC; 1.
DR SMART; SM00091; PAS; 2.
DR PROSITE; PS50888; HLH; 1.
DR PROSITE; PS50112; PAS; 2.
DR PROSITE; PS50112; PAS; 2.
SQ SEQUENCE 572 AA; 64162 MW; E4056207B02D1FD1 CRC64;

Query Match 87.8%; Score 36; DB 2; Length 572;

Best Local Similarity 85.7%; Pred. No. 98; Indels 0; Gaps 0;
Matches 6; Conservative 1; Mismatches 0;

QY 1 GNNYRPS 7

||:||||

Db 105 GSNYRPS 111

RESULT 14

ID Q96J63 PRELIMINARY; PRT; 578 AA.
AC Q96J63;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-MAR-2004 (TREMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE ARNTL2 protein (Fragment).
GN Name=ARNTL2;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1] SEQUENCE FROM N.A.
RP TISSUE=Placenta;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.C., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skaleka U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RC SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RA Strausberg R.;
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
 DR EMBL; BC000172; AA000172.3; --
 DR HSSP; Q9814; 1P97
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0004871; F:signal transducer activity; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR GO; GO:0007165; P:signal transduction; IEA.
 DR InterPro; IPR001092; HLH basic.
 DR InterPro; IPR001087; NucTranslocat.
 DR Pfam; PF00010; HLH; 1.
 DR Pfam; PF00989; PAS; 2.
 DR PRINTS; PR00785; NCTRNLOCATR.
 DR SMART; SM00353; HLH; 1.
 DR SMART; SM00091; PAS; 2.
 DR PROSITE; PS00888; HLH; 1.
 DR PROSITE; PS0112; PAS; 2.
 FT NON TER 1
 SQ SEQUENCE 578 AA; 280AF59AC0B082FD CRC64;

Query Match 87.8%; Score 36; DB 2; Length 578;
 Best Local Similarity 85.7%; Pred. No. 99;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
 DB 111 GSNYRPS 117
 |:|||||

RESULT 15

Q9NYQ5
 ID Q9NYQ5 PRELIMINARY; PRT; 585 AA.
 AC Q9NYQ5;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE BHLH-PAS transcription factor MOP9.
 GN Name=MOP9;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20325577; PubMed=10864977;
 RA Hogenesch J.B., Gu Y.Z., Moran S.M., Shimomura K., Radcliffe L.A.,
 RA Takahashi J.S., Bradfield C.A.;
 RT "The basic helix-loop-helix-PAS protein MOP9 is a brain-specific
 RT heterodimeric partner of circadian and hypoxia factors.";
 RL J. Neurosci. 20:RC83-RC83(2000).

CC -1- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
 DR EMBL; AF231338; AAF71306.1; --
 DR HSSP; Q99814; 1P97.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0004871; F:signal transducer activity; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR GO; GO:0007165; P:signal transduction; IEA.
 DR InterPro; IPR001092; HLH basic.
 DR InterPro; IPR001087; NucTranslocat.
 DR InterPro; IPR001610; PAS.
 DR Pfam; PF00010; HLH; 1.
 DR Pfam; PF00989; PAS; 2.
 DR PRINTS; PR00785; NCTRNLOCATR.
 DR SMART; SM00353; HLH; 1.
 DR SMART; SM00086; PAC; 1.
 DR SMART; SM00091; PAS; 2.
 DR PROSITE; PS00888; HLH; 1.
 DR PROSITE; PS0112; PAS; 2.
 SQ SEQUENCE 585 AA; 65556 MW; F7F44D78BB8ADF15 CRC64;

Query Match 87.8%; Score 36; DB 2; Length 585;
 Best Local Similarity 85.7%; Pred. No. 1e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNYRPS 7
 DB 118 GSNYRPS 124
 |:|||||

Search completed: April 28, 2005, 18:24:18
 Job time : 48.2812 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 27, 2004, 12:00:18 ; Search time 120 Seconds
(without alignments)
25.900 Million cell updates/sec

Title: US-10-088-639a-2_COPY_88_98
Perfect score: 65
Sequence: 1 NSWSSGTHPV 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database : A_Geneseq_29Jan04: *
1: geneseqp1980s: *
2: geneseqp1990s: *
3: geneseqp2000s: *
4: geneseqp2001s: *
5: geneseqp2002s: *
6: geneseqp2003as: *
7: geneseqp2003bs: *
8: geneseqp2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	65	100.0	249	4 AAB68087	Aab68087 An anti-a
2	51	78.5	11	5 AAO15005	Aao15005 GPI-link
3	49	75.4	110	6 ABR54947	Abr54947 IGG light
4	46	70.8	249	5 ABP45104	Abp45104 Human Bly
5	44	67.7	247	5 ABP45105	Abp45105 Human Bly
6	43	66.2	508	5 ABB06006	Abb06006 Human can
7	41	63.1	246	5 ABP45969	Abp45969 Human Bly
8	40	61.5	13	5 AAU70345	Aau70345 Human lam
9	40	61.5	107	6 ADB09502	Adb09502 Alloiococ
10	40	61.5	110	2 AAR15538	Aar15538 Anti-TGF
11	40	61.5	111	2 AAR12264	Aar12264 Anti-huma
12	40	61.5	112	6 ABR82488	Abr82488 Murine T
13	40	61.5	120	6 ABR82487	Abr82487 Murine T
14	40	61.5	249	5 ABP44529	Abp44529 Human Bly
15	40	61.5	249	5 ABP44342	Abp44342 Human Bly
16	40	61.5	469	5 ABP65370	Abp65370 Bifidobac
17	39	60.0	61	4 ABG01449	Abg01449 Novel hum
18	39	60.0	162	4 AAU28229	Aau28229 Novel hum
19	39	60.0	305	5 ABJ75568	Aau75568 Human den
20	39	60.0	305	6 ABJ39144	Abj39144 Molecule
21	39	60.0	311	4 AAB71910	Aab71910 Human ISO
22	39	60.0	324	4 AAU28081	Aau28081 Novel hum
23	39	60.0	683	6 ABU03132	Aau03132 Alpha amy
24	39	60.0	1336	5 AAU92973	Aau92973 Arabidops
25	39	60.0	1336	6 ADA15573	Ada15573 A. thalia

ADD30544	Plant yie
ADE31525	Plant yie
ABP01278	Human ORF
Aar72861	Fibronect
ABP56515	Human ant
AAO10624	Human pol
AAU4761	Propionib
ABM41280	Propionib
ABP44917	Human Bly
ABP44916	Human Bly
ABM65332	Propionib
AAM79385	Human pro
AAM41315	Human pol
AAM78401	Human pro
AAG29340	Arabidops
AAG29339	Arabidops
AAG29339	Human sec
ABP25919	Streptoco
ADB09266	Alloiococ
Aaw10210	Mature en
Aaw10209	Full leng
ABU24048	Protein e
ADA06278	Human cel
AAM25875	Human pro
AAM79159	Human pro
ABR58629	Human can
AAM80143	Human pro
ABG30355	Novel hum
ABO4874	LDL recep
AAG62970	Complemen
AAG62974	Complemen
AAG62984	Complemen
AAG62976	Complemen
AAG62980	Complemen
ABG78264	Human Fv
ABG31955	Human ant
AAG80194	Human gp9
ABBS1108	Human sec
ABO45365	Novel hum
ABO26845	Protein a
ABO4878	Proteinib
AAG6904	Proteinib
ABM61397	Proteinib
ABM63423	Proteinib
AAR80089	Human der
AAW95487	Human der
AAG60140	Proteinib
ABM56659	Proteinib
AAU47572	Proteinib
ABM44091	Proteinib
AAR22572	Light cha
AAV60148	Human end
AAG62943	Amino aci
AAG62969	Amino aci
AAV44616	Human ant
AAG62937	Amino aci
AAG62933	Amino aci
AAG62939	Amino aci
AAG80198	Human gp9
AAU19884	CEA-speci
AAU02509	Anti-adip
AAU02504	Anti-adip
AAU02513	Anti-adip
ABG78132	Human Fv
ABG91823	Human ant
AAG62947	Amino aci
ABBS1107	Human sec
ABO45364	Novel hum
ABO26844	Protein a
AAY29654	Mouse end
AAG8017	Human ace
AAW49694	Human ScF
AAO30914	DI-NHS76

99 37 56.9 233 5 ABG78152 Human Fv
 100 37 56.9 233 5 ABG91843 Human ant
 101 37 56.9 236 2 Aaw49690 Human SCF
 102 37 56.9 237 2 Aaw49691 Human SCF
 103 37 56.9 237 5 ABP46028 Human Bly
 104 37 56.9 237 5 ABP46101 Human Bly
 105 37 56.9 237 5 ABP45930 Human Bly
 106 37 56.9 238 5 ABP45920 Human Bly
 107 37 56.9 238 5 ABP46013 Human Bly
 108 37 56.9 238 5 ABP45896 Human Bly
 109 37 56.9 239 2 Aaw49692 Human SCF
 110 37 56.9 239 5 ABP46007 Human Bly
 111 37 56.9 239 5 ABP46024 Human Bly
 112 37 56.9 239 5 ABP46004 Human Bly
 113 37 56.9 239 5 ABP46027 Human Bly
 114 37 56.9 239 5 ABP44926 Human Bly
 115 37 56.9 240 5 ABP45380 Human Bly
 116 37 56.9 240 5 ABP45901 Human Bly
 117 37 56.9 240 5 ABP46002 Human Bly
 118 37 56.9 240 6 Aao311146 Human CM0
 119 37 56.9 241 5 ABP45900 Human Bly
 120 37 56.9 241 5 ABP46020 Human Bly
 121 37 56.9 241 5 ABP46049 Human Bly
 122 37 56.9 241 5 ABP45292 Human Bly
 123 37 56.9 242 5 ABP45667 Human Bly
 124 37 56.9 242 5 ABP45751 Human Bly
 125 37 56.9 242 5 ABP46035 Human Bly
 126 37 56.9 242 5 ABP45433 Human Bly
 127 37 56.9 242 5 ABP46010 Human Bly
 128 37 56.9 242 5 ABP45413 Human Bly
 129 37 56.9 242 5 ABP46095 Human Bly
 130 37 56.9 242 5 ABP45731 Human Bly
 131 37 56.9 243 5 ABP46045 Human Bly
 132 37 56.9 243 5 ABP45929 Human Bly
 133 37 56.9 243 5 ABP46003 Human Bly
 134 37 56.9 243 5 ABP45931 Human Bly
 135 37 56.9 243 5 ABP46046 Human Bly
 136 37 56.9 243 5 ABP45933 Human Bly
 137 37 56.9 243 5 ABP46096 Human Bly
 138 37 56.9 243 5 ABP45918 Human Bly
 139 37 56.9 244 5 ABP45922 Human Bly
 140 37 56.9 244 5 ABP45683 Human Bly
 141 37 56.9 244 5 ABP45681 Human Bly
 142 37 56.9 244 5 ABP46047 Human Bly
 143 37 56.9 244 6 Aao311139 Human CM0
 144 37 56.9 244 6 Aao311136 Human CM0
 145 37 56.9 245 5 ABP45908 Human Bly
 146 37 56.9 245 5 ABP45604 Human Bly
 147 37 56.9 245 5 ABP45815 Human Bly
 148 37 56.9 245 6 Aao311137 Human CM0
 149 37 56.9 246 5 ABG78329 Human Fv
 150 37 56.9 246 5 ABP45807 Human Bly

ALIGNMENTS

RESULT 1
 AAB68087
 ID AAB68087 standard; protein; 249 AA.

XX AC AAB68087;

XX XX
 DT 09-JUL-2001 (first entry)

XX DE An anti-alpha6beta4 integrin light chain linked to a heavy chain.

XX KW Gastrointestinal epithelial tumour cell; alpha6beta4 integrin;
 KW tumour-associated antigen; metastatic disease; malignant disease;
 XX tumour typing; tumour screening; tumour.

XX OS Synthetic.

OS Macaca fascicularis.

XX FH Location/Qualifiers
 Key 23. .33
 Region /note= "Complementarity determining region (CDR) 1 of the light chain"
 FT
 FT
 FT Region 49. .55
 FT /note= "Complementarity determining region (CDR) 2 of the light chain"
 FT
 FT Region 88. .98
 FT /note= "Complementarity determining region (CDR) 3 of the light chain"
 FT
 FT Peptide 110. .127
 FT /note= "linker"
 FT
 FT Region 158. .162
 FT /note= "Complementarity determining region (CDR) 1 of the heavy chain"
 FT
 FT Region 177. .193
 FT /note= "Complementarity determining region (CDR) 1 of the heavy chain"
 FT
 FT Region 226. .238
 FT /note= "Complementarity determining region (CDR) 1 of the heavy chain"
 FT
 WO200130854-A2.

03-MAY-2001.

26-OCT-2000; 2000WO-SE002082.

28-OCT-1999; 99SE-00003895.

(ACTI-) ACTIVE BIOTECH AB.

Brodin TN, Karlstroem PJ, Ohlsson LG, Tordsson MJ, Kearney PP;
 Nilsson BHK;

WPI; 2001-308619/32.

N-PSDB; AAF84797.

Novel antibody for diagnosis, treatment of human metastatic and malignant diseases, has binding structure for target structure displayed on cell surface of human gastrointestinal epithelial tumor and normal cells.

Claim 1; Page 55-56; 75pp; English.

The present sequence represents a Monkey antibody light chain linked to a heavy chain. The antibody binds to a target structure displayed in and on the cell surface of human gastrointestinal epithelial tumour cells and in a subpopulation of normal human gastrointestinal epithelial cells. The target structure especially comprises alpha6beta4 integrin. This is a tumour-associated antigen. The antibody, and its fragments, are useful for treating conditions based on anti-angiogenic mechanism in humans. They are useful for treating human metastatic and malignant disease, for in vitro, in vivo diagnosis and prognosis of human malignant disease, comprising tumour typing, tumour screening, tumour diagnosis and prognosis and monitoring premalignant conditions. Quantitative in vivo diagnosis is carried out by determining the localization of antibody to tumour deposits in humans

SQ Sequence 249 AA;

Query Match 100.0%; Score 65; DB 4; Length 249;

Best Local Similarity 100.0%; Pred. No. 0.011;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSWDSSGTHPV 11

DB 88 NSWDSSGTHPV 98

RESULT 2

AAO15005

ID AAO15005 standard; peptide; 11 AA.

XX AAO15005;
 AC
 XX 16-AUG-2002 (first entry)
 DT
 XX GPI-linked cell surface receptor-specific VL CDR3 peptide sequence.
 DE
 XX Ribosome display system; specific binding pair; sbp; anti-FITC; VL CDR3;
 KW scFv antibody generation; GPI-linked cell surface receptor.
 XX
 XX Unidentified.
 OS
 XX WO200175097-A2.
 PN
 XX 11-OCT-2001.
 PD
 XX 26-MAR-2001; 2001WO-GB001319.
 PF
 XX 31-MAR-2000; 2000US-0193802P.
 PR
 XX (CAMP-) CAMBRIDGE ANTIBODY TECHNOLOGY.
 PA
 XX Osbourn J, Holet T;
 PI
 XX WPI; 2002-010788/01.
 DR
 XX Obtaining specific binding pair member (I) that binds a complementary
 PT specific binding pair member of encapsulating specific binding pair
 PT member/ribosome complexes in a viral coat, comprises the use of a
 PT ribosome display system.
 PS
 XX Example 7; Page 42; 61pp; English.
 XX The invention comprises a method (ribosome display system) for obtaining
 CC a specific binding pair (sbp) member that binds a complementary sbp
 CC member of interest. The method involves incubating mRNA molecules
 CC encoding an sbp and lacking an in-frame stop codon, allowing ribosome
 CC translation of the mRNA to produce the encoded sbp member, forming
 CC complexes comprising ribosome, mRNA, and encoded sbp member displayed on
 CC the ribosome. The ribosome display system is useful for the selection of
 CC an sbp member able to bind a complementary sbp member. The present amino
 CC acid sequence represents a GPI-linked cell surface receptor-specific VL
 CC CDR3 peptide
 XX Sequence 11 AA;
 SQ
 Query Match 78.5%; Score 51; DB 5; Length 11;
 Best Local Similarity 81.8%; Pred. No. 0.077;
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 NSWDSGTHPV 11
 DB 1 NSWDSGNHVV 11
 RESULT 3
 ABR54947
 ID ABR54947 standard; protein; 110 AA.
 XX
 AC ABR54947;
 XX
 DT 30-JUN-2003 (first entry)
 XX
 DE IgG light chain clone HBL4a 3D6 SEQ ID NO:173.
 XX
 KW Engineered template; single primer amplification; antibody library;
 KW nucleic acid amplification.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 PN WO2003025202-A2.
 XX

PD 27-MAR-2003.
 XX
 PF 19-SEP-2002; 2002WO-US029889.
 XX
 PR 19-SEP-2001; 2001US-0323455P.
 XX
 PA (ALEX-) ALEXION PHARM INC.
 XX
 PI Bowdish KS, Frederickson S, Maruyama T, Lin Y, Renshaw M;
 XX
 DR WPI; 2003-313359/30.
 XX
 XX Amplifying nucleic acid by contacting engineered nucleic acid strand
 PT having predetermined sequence at one end and sequence complementary to
 PT predetermined sequence at other end, with primer having predetermined
 PT sequence.
 XX
 PS Example 3; Fig 8e; 68pp; English.
 XX
 CC The present invention describes a method (M1) for amplifying a nucleic
 CC acid strand. M1 comprises providing an engineered nucleic acid strand (S)
 CC having a predetermined sequence at one end and a sequence complementary
 CC to the predetermined sequence at the other, and contacting (S) with a
 CC primer having the predetermined sequence in the presence of a polymerase
 CC and nucleotides under conditions suitable for polymerisation of the
 CC nucleotides. Also described is an engineered nucleic acid strand (I)
 CC having a predetermined sequence at one end and a sequence complementary
 CC to the predetermined sequence at the other end. M1 is useful for
 CC amplifying a nucleic acid. M1 can be used for producing an antibody
 CC library. M1 is useful for preparing amplified products that can be
 CC ligated into a suitable expression vector, where the vector can be used
 CC to transform an appropriate host organism to produce the polypeptide or
 CC protein encoded by the target sequence. M1 is useful to amplify a family
 CC of related sequences to build a complex library such as, for example an
 CC antibody library. M1 is useful not only for producing large amounts of
 CC one target nucleic acid sequence, but also for amplifying simultaneously
 CC more than one different target nucleic acid sequence located on the same
 CC or different nucleic acid molecules. ACC52635 to ACC62753 and ABR54841 to
 CC ABR54998 represent sequence used in the exemplification of the present
 CC invention
 XX
 XX Sequence 110 AA;
 SQ
 Query Match 75.4%; Score 49; DB 6; Length 110;
 Best Local Similarity 88.9%; Pred. No. 2.1;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 3 WDSSGTHPV 11
 DB 92 WDSSGDHPV 100
 RESULT 4
 ABR45104
 ID ABR45104 standard; protein; 249 AA.
 XX
 AC ABR45104;
 XX
 DT 19-AUG-2002 (first entry)
 XX
 DE Human Blys binding scFv SEQ ID 1115.
 XX
 KW Blys; B lymphocyte stimulator; TNF superfamily; human; cytostatic;
 KW tumour necrosis factor; B cell proliferation; B cell differentiation;
 KW immunosuppressive; immunostimulant; immunomodulatory; antirheumatic;
 KW antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;
 KW systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;
 KW common variable immunodeficiency; acquired immunodeficiency syndrome.
 XX
 OS Homo sapiens.
 OS
 XX WO200202641-A1.
 PN
 XX

PD 10-JAN-2002.
 XX 15-JUN-2001; 2001WO-US019110.
 XX 16-JUN-2000; 2000US-0212210P.
 PR 17-OCT-2000; 2000US-0240816P.
 PR 16-MAR-2001; 2001US-0276248P.
 PR 21-MAR-2001; 2001US-0277379P.
 PR 25-MAY-2001; 2001US-0293499P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA (CAMB-) CAMBRIDGE ANTIBODY TECHNOLOGY.
 PI Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;
 XX WPI; 2002-114799/15.
 XX Antibodies against B Lymphocyte Stimulating polypeptides, useful for the
 PT diagnosis and treatment of cancers and immune disorders.
 XX Claim 1; Page 1733-1734; 3148pp; English.
 XX This invention describes novel antibodies that immunospecifically bind to
 CC B Lymphocyte Stimulator (BLYS) polypeptides. BLYS is a member of the
 CC tumour necrosis factor (TNF) super family and induces B cell
 CC proliferation and differentiation. The antibodies of the invention have
 CC cytostatic, immunosuppressive, immunostimulant, immunomodulatory,
 CC antirheumatic and antiAIDS activity and can be used in vaccines to
 CC inhibit the expression and activity of BLYS. The antibodies bind to BLYS
 CC and so may be used to detect and quantitate the presence of BLYS in
 CC biological samples and may be used in this way to diagnose disease
 CC associated with aberrant expression of BLYS. They may also be
 CC administered to treat diseases associated with aberrant BLYS expression
 CC and activity such as cancer, immune, and autoimmune disorders and
 CC diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis,
 CC immunodeficiency (e.g. common variable immunodeficiency (CVID) and
 CC acquired immunodeficiency syndrome (AIDS)). ABP43990-ABP47228 represent
 CC the antibodies and fragments of the antibodies described in the method of
 CC the invention
 XX SQ Sequence 249 AA;
 Query Match 70.8%; Score 46; DB 5; Length 249;
 Best Local Similarity 72.7%; Pred. No. 16;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 NSWDSSTHPV 11
 :|||||
 Db 228 HSWDSGSHV 238
 RESULT 5
 ID ABP45105
 ABP45105 standard; protein; 247 AA.
 XX AC ABP45105;
 XX 19-AUG-2002 (first entry)
 DT Human BLYS binding scFv SEQ ID 1116.
 XX BLYS; B lymphocyte stimulator; TNF superfamily; human; cytostatic;
 KW tumour necrosis factor; B cell proliferation; B cell differentiation;
 KW immunosuppressive; immunostimulant; immunomodulatory; antirheumatic;
 KW antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;
 KW systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;
 KW common variable immunodeficiency; acquired immunodeficiency syndrome.
 XX OS Homo sapiens.
 XX WO200202641-A1.
 XX 10-JAN-2002.

XX 15-JUN-2001; 2001WO-US019110.
 XX 16-JUN-2000; 2000US-0212210P.
 PR 17-OCT-2000; 2000US-0240816P.
 PR 16-MAR-2001; 2001US-0276248P.
 PR 21-MAR-2001; 2001US-0277379P.
 PR 25-MAY-2001; 2001US-0293499P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA (CAMB-) CAMBRIDGE ANTIBODY TECHNOLOGY.
 PI Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;
 XX WPI; 2002-114799/15.
 XX Antibodies against B Lymphocyte Stimulating polypeptides, useful for the
 PT diagnosis and treatment of cancers and immune disorders.
 XX Claim 1; Page 1734-1735; 3148pp; English.
 XX This invention describes novel antibodies that immunospecifically bind to
 CC B Lymphocyte Stimulator (BLYS) polypeptides. BLYS is a member of the
 CC tumour necrosis factor (TNF) super family and induces B cell
 CC proliferation and differentiation. The antibodies of the invention have
 CC cytostatic, immunosuppressive, immunostimulant, immunomodulatory,
 CC antirheumatic and antiAIDS activity and can be used in vaccines to
 CC inhibit the expression and activity of BLYS. The antibodies bind to BLYS
 CC and so may be used to detect and quantitate the presence of BLYS in
 CC biological samples and may be used in this way to diagnose disease
 CC associated with aberrant expression of BLYS. They may also be
 CC administered to treat diseases associated with aberrant BLYS expression
 CC and activity such as cancer, immune, and autoimmune disorders and
 CC diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis,
 CC immunodeficiency (e.g. common variable immunodeficiency (CVID) and
 CC acquired immunodeficiency syndrome (AIDS)). ABP43990-ABP47228 represent
 CC the antibodies and fragments of the antibodies described in the method of
 CC the invention
 XX SQ Sequence 247 AA;
 Query Match 67.7%; Score 44; DB 5; Length 247;
 Best Local Similarity 77.8%; Pred. No. 34;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 NSWDSSTGTH 9
 :|||||
 Db 226 HSWDSGSH 234
 RESULT 6
 ID ABB06006
 ABB06006 standard; protein; 508 AA.
 XX AC ABB06006;
 XX 09-MAY-2002 (first entry)
 DT Human cancer suppressing protein PP844 SRQ ID NO:26.
 XX Human; cancer suppression; cancer.
 XX OS Homo sapiens.
 XX CN131318-A.
 XX 19-SEP-2001.
 PD 14-MAR-2000; 2000CN-00111997.
 PF 14-MAR-2000; 2000CN-00111997.
 PR (SHAN-) SHANGHAI INST ONCOLOGY.

XX Gu J, Yang S;
 XX WPI; 2002-042196/06.
 DR N-PSDB; ABL39592.
 XX New human protein able to suppress growth of cancer cells and its
 PT encoding polynucleotide.
 XX Claim 1; Page 39 (Disclosure); 65pp; Chinese.
 XX The present invention describes human proteins with cancer suppressing
 CC activity. Also described are the polynucleotides encoding the proteins
 CC and a process for preparing the proteins by DNA recombination. The
 CC proteins and polynucleotides can be used in the treatment of diseases
 CC such as cancer. The present sequence represents a human cancer
 CC suppressing protein from the present invention
 XX SQ Sequence 508 AA;
 Query Match 66.2%; Score 43; DB 5; Length 508;
 Best Local Similarity 70.0%; Pred. No. 1.1e+02;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 NSWSSSGTHP 10
 |||||
 Db 160 NSWSSSRHP 169
 RESULT 7
 ABP45969
 ID ABP45969 standard; protein; 246 AA.
 XX AC ABP45969;
 XX DT 19-AUG-2002 (first entry)
 XX DE Human BlyS binding scFv SEQ ID 1980.
 XX BlyS; B lymphocyte stimulator; TNF superfamily; human; cytostatic;
 KW tumour necrosis factor; B cell proliferation; B cell differentiation;
 KW immunosuppressive; immunostimulant; immunomodulatory; antirheumatic;
 KW antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;
 KW systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;
 KW common variable immunodeficiency; acquired immunodeficiency syndrome.
 XX OS Homo sapiens.
 XX PN WO200202641-A1.
 XX PD 10-JAN-2002.
 XX PF 15-JUN-2001; 2001WO-US019110.
 XX PR 16-JUN-2000; 2000US-0212210P.
 PR 17-OCT-2000; 2000US-0240816P.
 PR 16-MAR-2001; 2001US-0276248P.
 PR 21-MAR-2001; 2001US-0277379P.
 PR 25-MAY-2001; 2001US-0293499P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA (CAMB-) CAMBRIDGE ANTIBODY TECHNOLOGY.
 XX Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;
 XX WPI; 2002-114799/15.
 XX Antibodies against B Lymphocyte Stimulating polypeptides, useful for the
 PT diagnosis and treatment of cancers and immune disorders.
 XX Claim 1; Page 2763-2764; 3148pp; English.
 XX This invention describes novel antibodies that immunospecifically bind to

CC B Lymphocyte Stimulator (BlyS) polypeptides. BlyS is a member of the
 CC tumour necrosis factor (TNF) super family and induces B cell
 CC proliferation and differentiation. The antibodies of the invention have
 CC cytostatic, immunosuppressive, immunostimulant, immunomodulatory,
 CC antirheumatic and antiAIDS activity and can be used in vaccines to
 CC inhibit the expression and activity of BlyS. The antibodies bind to BlyS
 CC and so may be used to detect and quantitate the presence of BlyS in
 CC biological samples and may be used in this way to diagnose disease
 CC associated with aberrant expression of BlyS. They may also be
 CC administered to treat diseases associated with aberrant BlyS expression
 CC and activity such as cancer, immune, and autoimmune disorders and
 CC diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis, and
 CC immunodeficiency (e.g. common variable immunodeficiency (CVID) and
 CC acquired immunodeficiency syndrome (AIDS)). ABP43990-ABP47228 represent
 CC the antibodies and fragments of the antibodies described in the method of
 CC the invention
 XX SQ Sequence 246 AA;
 Query Match 63.1%; Score 41; DB 5; Length 246;
 Best Local Similarity 81.8%; Pred. No. 1.1e+02;
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 NSWSSSGTHPV 11
 |||||
 Db 225 NSRDSSGTHLV 235
 RESULT 8
 AAU70345
 ID AAU70345 standard; peptide; 13 AA.
 XX AC AAU70345;
 XX DT 14-FEB-2002 (first entry)
 XX DE Human lambda III light chain CDR3.
 KW Immunoglobulin; antibody; light chain; heavy chain; CDR; FR;
 KW complementarity determining region; framework region; IgBP;
 KW transgenic plant; immunoglobulin binding protein array; IgM; IgG; IgA;
 KW IgD; IgE; IgY; IgM; kappa; lambda; CHBP.
 XX OS Homo sapiens.
 XX PN WO200183806-A1.
 XX PD 08-NOV-2001.
 XX PF 02-MAY-2001; 2001WO-US014349.
 XX PR 02-MAY-2000; 2000US-00563222.
 XX PA (EPIC-) EPICYTE PHARM INC.
 XX Hiatt AC, Hein MB;
 XX WPI; 2002-055482/07.
 XX Preparing immunoglobulin binding protein array in plant cells by
 PT transforming the cells with different polynucleotides encoding binding
 PT protein polypeptides specific to ligand, selecting plant cells for
 PT preparing array.
 XX Disclosure; Page 14; 129pp; English.
 XX The invention relates to transforming a population of cells (e.g. plant
 CC cells), comprising using a library of two different polynucleotides
 CC encoding different immunoglobulin binding protein (IgBP) polypeptides
 CC that specifically bind to a ligand or form one or more disulphide bonds
 CC with polypeptides in transfected cells, to generate an IgBP that binds to
 CC a ligand, and transformed plant cells are selected, and preparing an IgBP
 CC array in plant cells. At least one peptide sequence has at least 75%

CC sequence identity to a framework region (FR) of a native IgM, IgG, IgA, IgD, IgE, IgY, kappa or lambda immunoglobulin molecule. The method is useful for preparing an immunoglobulin binding protein array, preferably heavy chain binding protein (CHBP) array in eukaryotic cells especially plant cells (e.g. *Agrobacterium tumefaciens* or maize) or other eukaryotic cells (e.g. insect cells or mammalian cells). The CHBP is useful for discovery of e.g. screening assays of IgBPs having desired characteristics. The present sequence is a mammalian immunoglobulin derived peptide that may be incorporated into an IgBP of the invention

XX
XX
SQ Sequence 13 AA;

Query Match 61.5%; Score 40; DB 5; Length 13;
Best Local Similarity 66.7%; Pred. No. 6.4;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 SWDSSGTHP 10
Db 2 AWDSSDHP 10

RESULT 9
ADB09502
ID ADB09502 standard; protein; 107 AA.

XX ADB09502;
XX
XX 20-NOV-2003 (first entry)
XX
XX Alloiooccus otitis antigenic protein SEQ ID NO:3442.

DE Alloiooccus otitis; antigenic protein; immunogenic; immunisation;
KW gene therapy; Gram-positive bacterium; infection.

XX Alloiooccus otitis.

XX WO2003048304-A2.

XX 12-JUN-2003.

XX 25-NOV-2002; 2002WO-US036123.

XX 29-NOV-2001; 2001US-033377P.

XX 18-NOV-2002; 2002US-0426742P.

XX (AMHP) WYETH HOLDINGS CORP.

XX Fletcher LD, McMichael JC, Russell DP, Zagursky RJ;

XX WPI; 2003-505284/47.

XX N-PSDB; ADB09501.

XX New Alloiooccus otitis polynucleotides and polypeptides, useful for treating and diagnosing diseases, drug screening assays and monitoring of effects during drug clinical trials.

XX Claim 33; SEQ ID NO 3442; 1019pp; English.

XX The present invention describes an isolated polynucleotide (1) of
XX Alloiooccus otitis genomic DNA, which encodes an antigenic protein.
XX Alloiooccus otitis is a Gram-positive bacterium. Also described: (1)
XX an isolated polypeptide that is encoded by the polynucleotide (1); (2) an
XX expression vector comprising the novel isolated polynucleotide (1), its
XX complement, degenerate variant or fragment; (3) a genetically engineered
XX host cell, transfected, transformed or infected with the vector of (2);
XX (4) an antibody specific for the polypeptide of (1); (5) an immunogenic
XX composition comprising the polypeptide, its complement, biological
XX equivalent or fragment, or the polynucleotide that is comprised in the
XX expression vector; (6) a pharmaceutical composition comprising the
XX polypeptide of (1) and a carrier; (7) a protein chip comprising an array
XX of the polypeptides of (1), their biological equivalent or fragment; (8)
XX immunising against Alloiooccus otitis by administering to a host the
XX immunogenic composition; (9) detecting and/or identifying Alloiooccus

CC otitis in the biological sample; (10) a kit comprising a container
CC containing the novel polynucleotide, its degenerate variant or fragment,
CC or the antibody of (4); and (11) producing a polypeptide by culturing the
CC genetically engineered host cell under conditions suitable to produce the
CC polypeptide from the culture. (1) can be used in gene therapy. The
CC polynucleotides, polypeptides, antibodies and compositions of the present
CC invention can be used for treating and diagnosing diseases, drug
CC screening assays and monitoring of effects during drug clinical trials.
CC The polynucleotides are useful for expressing and detecting Alloiooccus
CC otitis. The present sequence represents an Alloiooccus otitis
XX antigen protein from the present invention.

SQ Sequence 107 AA;

Query Match 61.5%; Score 40; DB 6; Length 107;
Best Local Similarity 70.0%; Pred. No. 63;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 NSWDSSGTHP 10
Db 65 SSWSSSGTWP 74

RESULT 10
AAW15538
ID AAW15538 standard; protein; 110 AA.

XX AAW15538;

XX 27-NOV-1997 (first entry)

DE Anti-TGF beta-1 scFv antibody 10A6 VL domain.

XX Transforming growth factor beta-1; TGF-beta-1; human;

XX antibody engineering; scFv; phage display; lung fibrosis;

XX arterial injury; proliferative retinopathy; retinal detachment;

XX adult respiratory distress syndrome; liver cirrhosis;

XX post myocardial infarction; post-angioplasty restenosis; scleroderma;

XX vascular disease; cataract; glaucoma; scarring; glomerulonephritis;

XX osteoporosis; immune disease; inflammation; rheumatoid arthritis;

XX macrophage deficiency disease; macrophage pathogen infection; therapy.

XX Homo sapiens.

XX GB2305921-A.

XX 23-APR-1997.

XX 07-OCT-1996; 96GB-00020920.

XX 06-OCT-1995; 95GB-00020486.

XX 19-JAN-1996; 96GB-00001081.

XX (CAMB-) CAMBRIDGE ANTIBODY TECHNOLOGY.

XX Thompson JE, Vaughan TJ, Williams AJ, Green JA, Jackson RH;

XX Bacon L, Johnson KS, Wilton AJ, Tempest PR, Pope AR;

XX WPI; 1997-215360/20.

XX N-PSDB; AAT60384.

XX Agent contg. antigen-binding domain of human antibody to transforming
XX growth factor beta 1 or 2 - and nucleic acid encoding it, used to
XX neutralise effects of TGF, e.g. for control of fibrosis, immune and
XX inflammatory disease.

XX Claim 16; Fig 1b(ii); 184pp; English.

XX This polypeptide sequence comprises the VL domain of human scFv antibody
XX 10A6, which is specific for transforming growth factor (TGF) beta-1. It
XX is encoded by a gene (AAT60384) isolated from a peripheral blood
XX lymphocyte library. The antigen-binding domains of human antibodies (see
XX AAW15522-40) to TGF beta-1 and/or beta-2 can be used to counter the

CC adverse effects of TGF beta, such as (i) promotion of fibrosis (in
 CC dermal, ocular or keloid scarring, lung fibrosis, arterial injury,
 CC proliferative retinopathy, retinal detachment, adult respiratory distress
 CC syndrome, liver cirrhosis, post myocardial infarction, post-angioplasty
 CC restenosis, scleroderma, vascular disorders, cataract, glaucoma, or esp.
 CC neural scarring and glomerulonephritis, also (not claimed) osteoporosis),
 CC or (ii) immune and inflammatory diseases (e.g. rheumatoid arthritis,
 CC macrophage deficiency diseases or macrophage pathogen infection). Nucleic
 CC acids encoding human antibody VH and VL can be used for prodn. of
 CC recombinant antigen-binding domains. These are highly specific, have low
 CC dissociation constants (pref. less than 5 nM) and low IC50s for
 CC neutralisation
 XX
 SQ Sequence 110 AA;

Query Match 61.5%; Score 40; DB 2; Length 110;
 Best Local Similarity 88.9%; Pred. No. 65;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NSWDSSGTH 9
 || |||||
 Db 88 NSRDSSGTH 96

RESULT 11

AAR12264
 ID AAR12264 standard; protein; 111 AA.

XX AAR12264;

AC AAR12264;

XX 25-MAR-2003 (revised)

DT 15-AUG-1991 (first entry)

XX Anti-human Rhd PAG-1 Mab (VL chain).

DE Monoclonal antibody; rhesus D; blood-typing; CDR;

XX Haemolytic disease of the newborn; HDN.

KW Homo sapiens.

XX

OS

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

CC The chimaeric anti-Rhd antibodies can be used for diagnosis and therapy,
 CC and are capable of providing blood- typing reagents of high specificity
 CC and reliability. They can also be used in passive immunisation to prevent
 CC haemolytic disease of the newborn. (Updated on 25-MAR-2003 to correct PA
 CC field.) (Updated on 25-MAR-2003 to correct PI field.)
 XX

SQ Sequence 111 AA;

Query Match 61.5%; Score 40; DB 2; Length 111;
 Best Local Similarity 75.0%; Pred. No. 66;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 WDSSGTHP 10
 |||||
 Db 90 WDSSSAHP 97

RESULT 12

ABR82488

ID ABR82488 standard; protein; 112 AA.

XX ABR82488;

AC ABR82488;

XX 20-NOV-2003 (first entry)

DT Murine T cell receptor.Vgamma6 chain.

DE T cell receptor; TCR; gastrointestinal; antineoplastic; cytostatic;

XX tuberculostatic; dermatological; antibacterial; virucide; gynaecological;

XX cell therapy; Vgamma; mouse.

XX Mus musculus.

OS WO2003060097-A2.

XX 24-JUL-2003.

PD 10-JAN-2003; 2003WO-US000728.

XX 10-JAN-2002; 2002US-0347285P.

XX (NAJE-) NAT JEWISH MEDICAL & RES CENT.

XX O'Brien RL, Born WK, Roark CL, Aydinug MK;

XX WPI; 2003-598525/56.

XX N-PSDB; ACF35982.

XX Regulating a gammadelta T-cell mediated immune response in a mammal,

XX useful for treating inflammation in intestine, skin, lungs or

XX reproductive tract, comprises administering to the mammal a soluble

XX gammadelta T cell receptor.

XX Example; Page 58-59; 71pp; English.

XX The invention relates to regulating a gammadelta T-cell mediated immune

XX response in a mammal and involves administering to the mammal a soluble

XX gammadelta T cell receptor (TCR). The method is useful for treating

XX patients having, or are at risk of developing an intestinal condition,

XX e.g. Crohn's disease, ischaemic colitis, irritable bowel disease, and

XX colon cancer; a lung condition associated with inflammation such as

XX airway hyperresponsiveness, pneumonia, tuberculosis, and a primary or

XX metastatic lung tumour; a skin condition associated with inflammation or

XX such as skin lesion caused by bacterial infection, viral infection or

XX laceration, and a skin cancer; or a condition associated with

XX inflammation of the reproductive tract such as infection caused by

XX bacterial or viral infection that involve the epithelial mucosal lining,

XX a tubal infection, preventing tubal factor infertility, and a cancer

XX selected from ovarian, cervical, uterine, prostate or testicular cancers.

XX Sequences ABR82485-489 represent murine TCR Vgamma chains used in the

XX present invention

XX Sequence 112 AA;

Query Match 61.5%; Score 40; DB 6; Length 112;
 Best Local Similarity 77.8%; Pred. No. 67;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 WDSSGTHPV 11
 |||||
 Db 104 WDSSGFHKV 112

RESULT 13
 ABR82487
 ID ABR82487 standard; protein; 120 AA.
 XX ABR82487;
 AC ABR82487;
 XX ABR82487;
 DT 20-NOV-2003 (first entry)
 XX Murine T cell receptor Vgamma5 chain.
 DE Murine T cell receptor Vgamma5 chain.
 KW T cell receptor; TCR; gastrointestinal; antiinflammatory; cytostatic;
 KW tuberculosis; dermatological; antibacterial; virucide; gynaecological;
 KW cell therapy; Vgamma; mouse.
 XX cell therapy; Vgamma; mouse.
 OS Mus musculus.
 XX Mus musculus.
 PN WO2003060097-A2.
 XX WO2003060097-A2.
 PD 24-JUL-2003.
 XX 24-JUL-2003.
 PF 10-JAN-2003; 2003WO-US000728.
 XX 10-JAN-2003; 2003WO-US000728.
 PR 10-JAN-2002; 2002US-0347285P.
 XX 10-JAN-2002; 2002US-0347285P.
 PA (NAJE-) NAT JEWISH MEDICAL & RES CENT.
 XX (NAJE-) NAT JEWISH MEDICAL & RES CENT.
 PI O'Brien RL, Born WK, Roark CL, Aydtung MK;
 XX O'Brien RL, Born WK, Roark CL, Aydtung MK;
 DR WPI; 2003-598525/56.
 DR N-PSDB; ACF35981.
 XX WPI; 2003-598525/56.
 PT Regulating a gammadelta T-cell mediated immune response in a mammal,
 PT useful for treating inflammation in intestine, skin, lungs or
 PT reproductive tract, comprises administering to the mammal a soluble
 PT gammadelta T cell receptor.
 XX Regulating a gammadelta T-cell mediated immune response in a mammal,
 PS Example; Page 57-58; 71pp; English.
 XX Example; Page 57-58; 71pp; English.
 CC The invention relates to regulating a gammadelta T-cell mediated immune
 CC response in a mammal and involves administering to the mammal a soluble
 CC gammadelta T cell receptor (TCR). The method is useful for treating
 CC patients having, or are at risk of developing an intestinal condition,
 CC e.g. Crohn's disease, ischaemic colitis, irritable bowel disease, and
 CC colon cancer; a lung condition associated with inflammation such as
 CC airway hyperresponsiveness, pneumonia, tuberculosis, and a primary or
 CC metastatic lung tumour; a skin condition associated with inflammation
 CC such as skin lesion caused by bacterial infection, viral infection or
 CC laceration, and a skin cancer; or a condition associated with
 CC inflammation of the reproductive tract such as infection caused by
 CC bacterial or viral infection that involve the epithelial mucosal lining,
 CC a tubal infection, preventing tubal factor infertility, and a cancer
 CC selected from ovarian, cervical, uterine, prostate or testicular cancers.
 CC Sequences ABR82485-489 represent murine TCR Vgamma chains used in the
 CC present invention
 XX present invention
 SQ Sequence 120 AA;

Query Match 61.5%; Score 40; DB 6; Length 120;
 Best Local Similarity 77.8%; Pred. No. 72;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 WDSSGTHPV 11
 |||||
 Db 104 WDSSGFHKV 112

Db 112 WDSSGFHKV 120

RESULT 14
 ABP44529
 ID ABP44529 standard; protein; 249 AA.
 XX ABP44529;
 AC ABP44529;
 XX ABP44529;
 DT 19-AUG-2002 (first entry)
 XX Human BlyS binding scFv SEQ ID 540.
 DE Human BlyS binding scFv SEQ ID 540.
 KW BlyS; B lymphocyte stimulator; TNF superfamily; human; cytostatic;
 KW tumour necrosis factor; B cell proliferation; B cell differentiation;
 KW immunosuppressive; immunostimulant; immunomodulatory; antirheumatic;
 KW antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;
 KW systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;
 KW common variable immunodeficiency; acquired immunodeficiency syndrome.
 XX common variable immunodeficiency; acquired immunodeficiency syndrome.
 OS Homo sapiens.
 XX Homo sapiens.
 PN WO200202641-A1.
 XX WO200202641-A1.
 PD 10-JAN-2002.
 XX 10-JAN-2002.
 PF 15-JUN-2001; 2001WO-US019110.
 XX 15-JUN-2001; 2001WO-US019110.
 PR 16-JUN-2000; 2000US-0212210P.
 PR 17-OCT-2000; 2000US-0240816P.
 PR 16-MAR-2001; 2001US-0276248P.
 PR 21-MAR-2001; 2001US-0277379P.
 PR 25-MAY-2001; 2001US-0293499P.
 XX 25-MAY-2001; 2001US-0293499P.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PA (CAMB-) CAMBRIDGE ANTIBODY TECHNOLOGY.
 PI Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;
 XX Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;
 DR WPI; 2002-114799/15.
 XX WPI; 2002-114799/15.
 PT Antibodies against B Lymphocyte Stimulating polypeptides, useful for the
 PT diagnosis and treatment of cancers and immune disorders.
 XX Antibodies against B Lymphocyte Stimulating polypeptides, useful for the
 PS Claim 1; Page 1047-1048; 3148pp; English.
 XX Claim 1; Page 1047-1048; 3148pp; English.
 CC This invention describes novel antibodies that immunospecifically bind to
 CC B Lymphocyte Stimulator (BlyS) polypeptides. BlyS is a member of the
 CC tumour necrosis factor (TNF) super family and induces B cell
 CC proliferation and differentiation. The antibodies of the invention have
 CC cytostatic, immunosuppressive, immunostimulant, immunomodulatory,
 CC antirheumatic and antiAIDS activity and can be used in vaccines to
 CC inhibit the expression and activity of BlyS. The antibodies bind to BlyS
 CC and so may be used to detect and quantitate the presence of BlyS in
 CC biological samples and may be used in this way to diagnose disease
 CC associated with aberrant expression of BlyS. They may also be
 CC administered to treat diseases associated with aberrant BlyS expression
 CC and activity such as cancer, immune, and autoimmune disorders and
 CC diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis,
 CC immunodeficiency (e.g. common variable immunodeficiency (CVID) and
 CC acquired immunodeficiency syndrome (AIDS)). ABP43990-ABP47228 represent
 CC the antibodies and fragments of the antibodies described in the method of
 CC the invention
 XX the invention
 SQ Sequence 249 AA;

Query Match 61.5%; Score 40; DB 5; Length 249;
 Best Local Similarity 88.9%; Pred. No. 1,6e+02;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NSWDSSGTH 9
 |||||
 Db 228 NSRDSSGTH 236

RESULT 15
ABP44342
ID ABP44342 standard; protein; 249 AA.
XX
AC ABP44342;
XX
DT 19-AUG-2002 (first entry)
XX
DE Human BlyS binding scFv SEQ ID 353.
XX
KW BlyS; B lymphocyte stimulator; TNF superfamily; human; cytostatic;
KW tumour necrosis factor; B cell proliferation; B cell differentiation;
KW immunosuppressive; immunostimulant; immunomodulatory; antirheumatic;
KW antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;
KW systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;
KW common variable immunodeficiency; acquired immunodeficiency syndrome.
XX
OS Homo sapiens.
XX
PN WO200202641-A1.
XX
PD 10-JAN-2002.
XX
PF 15-JUN-2001; 2001WO-US019110.
XX
PR 16-JUN-2000; 2000US-0212210P.
PR 17-OCT-2000; 2000US-0240816P.
PR 16-MAR-2001; 2001US-0276248P.
PR 21-MAR-2001; 2001US-0277379P.
PR 25-MAY-2001; 2001US-0293499P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PA (CAMB-) CAMBRIDGE ANTIBODY TECHNOLOGY.
XX
PI Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;
XX
DR WPI; 2002-114799/15.
XX
PT Antibodies against B Lymphocyte Stimulating polypeptides, useful for the
PT diagnosis and treatment of cancers and immune disorders.
XX
PS Claim 1; Page 825-826; 3148pp; English.
XX
CC This invention describes novel antibodies that immunospecifically bind to
CC B Lymphocyte Stimulator (BlyS) polypeptides. BlyS is a member of the
CC tumour necrosis factor (TNF) super family and induces B cell
CC proliferation and differentiation. The antibodies of the invention have
CC cytostatic, immunosuppressive, immunostimulant, immunomodulatory,
CC antirheumatic and antiAIDS activity and can be used in vaccines to
CC inhibit the expression and activity of BlyS. The antibodies bind to BlyS
CC and so may be used to detect and quantitate the presence of BlyS in
CC biological samples and may be used in this way to diagnose disease
CC associated with aberrant expression of BlyS. They may also be
CC administered to treat diseases associated with aberrant BlyS expression
CC and activity such as cancer, immune, and autoimmune disorders and
CC diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis,
CC immunodeficiency (e.g. common variable immunodeficiency (CVID) and
CC acquired immunodeficiency syndrome (AIDS)). ABP4390-ABP47228 represent
CC the antibodies and fragments of the antibodies described in the method of
CC the invention
XX
SQ Sequence 249 AA;

Query Match 61.5%; Score 40; DB 5; Length 249;
Best Local Similarity 88.9%; Pred. No. 1.6e+02;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 NSWDSSGTH 9
DB 228 NSRDSGTH 236

RESULT 16
ABP65370
ID ABP65370 standard; protein; 469 AA.
XX
AC ABP65370;
XX
DT 19-NOV-2002 (first entry)
XX
DE Bifidobacterium longum NCC2705 ORF amino acid sequence SEQ ID NO:114.
XX
KW Bifidobacterium longum NCC2705; Bifidobacterium; bacterial;
KW antidiarrheic; antibacterial; inhibitor of Salmonella; detection;
KW identification; lactic acid bacterium; diarrhoea; pathogenic bacteria;
KW rotavirus; food composition; pharmaceutical composition.
XX
OS Bifidobacterium longum.
XX
PN EP1227152-A1.
XX
PD 31-JUL-2002.
XX
PF 30-JAN-2001; 2001EP-00102050.
XX
PR 30-JAN-2001; 2001EP-00102050.
XX
PA (NEST) SOC PROD NESTLE SA.
XX
DR WPI; 2002-668397/72.
XX
PT Novel polynucleotide comprising Bifidobacterium genome sequence useful as
PT a probe or primer for detecting and/or identifying Bifidobacterium longum
PT in a biological sample.
XX
PS Claim 3; SEQ ID NO 114; 80pp; English.
XX
CC The present invention describes a polynucleotide (I) comprising a
CC sequence of a Bifidobacterium genome selected from the nucleotide
CC sequences given in ABQ81842 and ABQ81843, or a sequence exhibiting at
CC least 90% identity or which hybridises with the sequences given in
CC ABQ81842 and ABQ81843. Also described is a polynucleotide (II) encoding a
CC fusion protein, comprising a sequence selected from 1097 sequences given
CC in ABP65258 to ABP65354 ligated in frame to a polynucleotide encoding a
CC heterologous polypeptide. (I) has antidiarrheic and antibacterial
CC activities, and can be used as an inhibitor of Salmonella. (I) (which is
CC a probe) is useful for the detection and/or identification of
CC Bifidobacterium longum in a biological sample. A carrier containing the
CC lactic acid bacterium Bifidobacterium longum NCC2705 (CNCM I-2618) can be
CC used for preventing and/or treating diarrhoea brought about by pathogenic
CC bacteria and/or rotavirus. The carrier is a food composition selected
CC from milk, yogurt, curd, cheese, fermented milks, milk based fermented
CC products, ice-creams, fermented cereal based products, milk based
CC powders, infant formula, pet food or a pharmaceutical composition
CC selected from tablets, liquid bacterial suspensions, dried oral
CC supplement, wet oral supplement, dry tube feeding or wet tube feeding.
CC (I) is useful in DNA arrays or chips to carry out analysis of the
CC expression of the Bifidobacterium genome. ABQ81844 to ABQ81850 represent
CC Bifidobacterium related nucleotide sequences given in the Sequence
CC Listing from the present invention but not mentioned further within the
CC specification. N.B. The sequence data for this patent is not represented
CC in the printed specification but is based on sequence information
CC supplied by the European Patent Office
XX
SQ Sequence 469 AA;

Query Match 61.5%; Score 40; DB 5; Length 469;
Best Local Similarity 87.5%; Pred. No. 3.2e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 NSWDSSTG 8
DB 19 NSWDSYGT 26

```

RESULT 17
ABG01449
ID ABG01449 standard; protein; 61 AA.
XX
XX
AC ABG01449;
XX
XX
DT 13-FEB-2002 (first entry)
XX
XX
DE Novel human diagnostic protein #1440.
XX
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX
OS Homo sapiens.
XX
XX
PN WO200175067-A2.
XX
XX
PD 11-OCT-2001.
XX
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
XX
PR 31-MAR-2000; 2000US-00540217.
XX
PR 23-AUG-2000; 2000US-00649167.
XX
XX
PA (HYSE-) HYSEQ INC.
XX
XX
PI Dmanac RT, Liu C, Tang YT;
XX
XX
DR WPI; 2001-639362/73.
XX
DR N-PSDB; AAS65636.
XX
XX
New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
XX
PS Claim 20; SEQ ID NO 31808; 103pp; English.
XX
XX
The invention relates to isolated polynucleotide (I) and polypeptide (II)
sequences. (I) is useful as hybridisation probes, polymerase chain
reaction (PCR) primers, oligomers, and for chromosome and gene mapping.
CC
CC
in and recombinant production of (II). The polynucleotides are also used
CC
CC
in diagnostics as expressed sequence tags for identifying expressed
CC
genes. (I) is useful in gene therapy techniques to restore normal
CC
activity of (II) or to treat disease states involving (II). (II) is
CC
useful for generating antibodies against it, detecting or quantitating a
CC
polypeptide in tissue, as molecular weight markers and as a food
CC
supplement. (II) and its binding partners are useful in medical imaging
CC
of sites expressing (II). (I) and (II) are useful for treating disorders
CC
involving aberrant protein expression or biological activity. The
CC
polypeptide and polynucleotide sequences have applications in
CC
diagnostics, forensics, gene mapping, identification of mutations
CC
responsible for genetic disorders or other traits to assess biodiversity
CC
and to produce other types of data and products dependent on DNA and
CC
amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
CC
amino acid sequences of the invention. Note: The sequence data for this
CC
patent did not appear in the printed specification, but was obtained in
CC
electronic format directly from WIPO at
CC
ftp.wipo.int/pub/published_pct_sequences
XX
XX
SQ Sequence 61 AA;
XX
Query Match 60.0%; Score 39; DB 4; Length 61;
Best Local Similarity 50.0%; Pred. No. 50;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 1 NSWDSSGTHP 10
: ||| |||
Db 10 HOWDEQGAHP 19
: ||| |||
RESULT 18

```

```

AAU28269
ID AAU28269 standard; protein; 162 AA.
XX
XX
AC AAU28269;
XX
XX
DT 18-DEC-2001 (first entry)
XX
XX
DE Novel human secretory protein, Seq ID No 626.
XX
XX
KW Human; secreted protein; arthritis; Crohn's disease; sepsis; shock;
KW ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;
KW transgenic animal; Alzheimer's disease; Parkinson's disease; burn;
KW amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;
KW ulcer; osteoporosis; bone degenerative disorder; periodontal disease;
KW gut protection; lung; liver fibrosis; immune deficiency; infection;
KW severe combined immunodeficiency; SCID; autoimmune disorder; allergy;
KW multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;
KW fertility; analgesic; pain; antigen.
XX
XX
OS Homo sapiens.
XX
XX
PN WO200166689-A2.
XX
XX
PD 13-SEP-2001.
XX
XX
PF 05-MAR-2001; 2001WO-US004942.
XX
XX
PR 07-MAR-2000; 2000US-00519705.
XX
PR 19-MAY-2000; 2000US-00574454.
XX
PR 17-JUN-2000; 2000US-00596193.
XX
PR 14-JUL-2000; 2000US-00616847.
XX
PR 19-SEP-2000; 2000US-00665363.
XX
PR 20-OCT-2000; 2000US-00693267.
XX
XX
PA (HYSE-) HYSEQ INC.
XX
XX
PI Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;
PI Zhao QA, Yang Y, Dmanac RT, Zhang J, Chen R, Xue AJ, Wang J;
XX
XX
WPI; 2001-589934/66.
XX
N-PSDB; AAS45169.
XX
Novel polypeptides and nucleic acids obtained from cDNA libraries
PT prepared from various human tissues, for diagnosis and treatment of
PT cancer, neurological, inflammatory, and autoimmune disorders.
XX
XX
Example 2; SEQ ID NO 626; 107pp; English.
XX
XX
The invention relates to novel isolated human secreted polypeptides (I)
and polynucleotides (II). (I) and (II) are useful for treating
CC
inflammatory conditions such as arthritis, sepsis, immune responses, and is
CC
ischaemia-reperfusion injury, shock, sepsis, immune responses, and is
CC
involved in increasing haematopoiesis, stem cell survival, bone growth
CC
and remodeling. (I), (II) and modulators of (II) are useful for
CC
prophylaxis or treatment of one or more cancers. (II) is also useful for
CC
creating transgenic animals useful for studying the in vivo activities of
CC
the polypeptide as well as for studying modulators of the polypeptides.
CC
(I) induces the proliferation of neural cells and regeneration of nerve
CC
and brain tissue and is useful for the treatment of central and
CC
peripheral nervous system diseases and neuropathies, such as Alzheimer's,
CC
Parkinson's disease, Huntington's disease, and amyotrophic lateral
CC
sclerosis. In addition, (I) is involved in chemotactic or chemokinetic
CC
activity, regulation of haematopoiesis and is useful for treating myeloid
CC
or lymphoid cell disorders, platelet disorders such as thrombocytopenia
CC
for regeneration of bone, cartilage, tendon, ligament and/or nerve
CC
tissue growth, and in tissue repair, healing of burns, incisions, ulcers,
CC
for treating osteoporosis, osteoarthritis, bone degenerative disorders,
CC
or periodontal disease. Furthermore, (I) is also useful for gut
CC
protection or regeneration and treatment of lung or liver fibrosis, and
CC
reperfusion injury in various tissues, various immune deficiencies and
CC
disorders including severe combined immunodeficiency (SCID), bacterial or
CC
fungal infections, autoimmune disorders e.g. multiple sclerosis,
CC
rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic

```

CC reactions and conditions, such as asthma or other respiratory problems.
 CC In addition, (I) affects biorhythms or circadian cycles of rhythms,
 CC fertility, metabolism, catabolism, anabolism, storage or elimination of
 CC dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides
 CC analgesic effects or other pain reducing effects, immunoglobulin like
 CC activity and can act as an antigen in a vaccine composition to raise an
 CC immune response. AAU28020-AAU28395 represent novel human secreted protein
 CC amino acid sequences of the invention
 XX
 SQ Sequence 162 AA;
 Query Match 60.0%; Score 39; DB 4; Length 162;
 Best Local Similarity 77.8%; Pred. No. 1.5e+02;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 2 SWDSSGTHP 10
 DB 67 SWDFSGTTP 75
 RESULT 19
 AAU75568
 ID AAU75568 standard; protein; 305 AA.
 XX
 AC AAU75568;
 DT 08-MAY-2002 (first entry)
 DE Human dendritic cell-derived protein #2.
 XX
 KW Human; dendritic cell; immune system; haematopoietic cell; immunogenic.
 XX
 OS Homo sapiens.
 PH Key Location/Qualifiers
 FT Misc-difference 44 /label= unknown
 FT /note= "Encoded by aaw"
 FT Misc-difference 64 /label= unknown
 FT /note= "Encoded by awt"
 XX
 PN EP1174502-A1.
 XX
 PD 23-JAN-2002.
 XX
 PF 18-JUL-2000; 2000EP-00306087.
 XX
 PR 18-JUL-2000; 2000EP-00306087.
 XX
 PA (SCHE) SCHERING CORP.
 XX
 PI Risoan M, Bridon J, Duhon T, Briere F, Bates E;
 XX
 XX WPI; 2002-156690/21.
 DR N-PSDB; ABK14747.
 XX
 PT Novel polypeptides and polynucleotides related to dendritic cells of the
 PT immune system, useful in the treatment of conditions associated with
 PT abnormal physiology or development.
 XX
 XX Claim 1; Page 21-22; 34pp; English.
 XX
 CC The invention relates to an isolated polypeptide (I) related to dendritic
 CC cells of the immune system and the polynucleotide (II) encoding (I). (I)
 CC is useful for screening for candidate therapeutic agents. (II) is useful
 CC for studying the presence, amount, distribution and normalcy of certain
 CC gene products produced by or expressed on dendritic cells. (II) is also
 CC useful to facilitate the discovery of compositions and methods useful for
 CC diagnosing and treating certain disease states. (I) and (II) are useful
 CC for diagnostic and therapeutic purposes. (I) is useful in binding
 CC studies, for construction and expression of modified molecules, for
 CC structure/function studies and for the preparation of polyclonal and

CC monoclonal antibodies. (I) is also useful as immunogenic components (e.g.
 CC antigens) for preparing antibodies or as targets for binding agent
 CC studies. (I) is useful in the treatment of conditions associated with
 CC abnormal physiology or development, e.g. disease or disorder associated
 CC with abnormal expression or abnormal signaling by a dendritic cell. (I)
 CC is also useful in the regulation and development of haematopoietic cells.
 CC The binding agent is useful to isolate and purify the immunogenic
 CC components by immunoaffinity chromatography, and as probes to distinguish
 CC tissue and cell type distribution. These are also useful to screen
 CC expression libraries for particular expression products. (II) is useful
 CC as templates for the recombinant production of peptides and polypeptides,
 CC as probes and primers for the detection of the human genes, for
 CC chromosome mapping and as probes or to design PCR primers to identify
 CC homologous genes in other mammalian species. The present sequence
 CC represents the amino acid sequence of human dendritic cell-derived
 CC protein #2
 XX
 SQ Sequence 305 AA;
 Query Match 60.0%; Score 39; DB 5; Length 305;
 Best Local Similarity 50.0%; Pred. No. 2.9e+02;
 Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 1 NSWSSGTHP 10
 DB 254 HQWDEQGAP 263
 RESULT 20
 ABJ39144
 ID ABJ39144 standard; protein; 305 AA.
 XX
 AC ABJ39144;
 DT 16-OCT-2003 (first entry)
 DE Molecule for disease detection and treatment (MDDT)-54 protein sequence.
 XX
 KW Human; Molecules for Disease Detection and Treatment; MDDT;
 KW MDDT-antagonist; MDDT-agonist; gene therapy; cardiac; cytostatic;
 KW neuroprotective; cardiovascular disorder; neurological disorder;
 KW cell proliferative disorder; autoimmune disorder; inflammatory disorder;
 KW developmental disorder; cancer; steroid metabolic response.
 XX
 OS Homo sapiens.
 XX
 PN WO2003052049-A2.
 XX
 PD 26-JUN-2003.
 XX
 PF 09-JUL-2002; 2002WO-US021767.
 XX
 PR 09-JUL-2001; 2001US-0304298P.
 PR 13-JUL-2001; 2001US-0305324P.
 PR 19-JUL-2001; 2001US-0307003P.
 PR 27-JUL-2001; 2001US-0308185P.
 PR 03-AUG-2001; 2001US-0310096P.
 PR 10-AUG-2001; 2001US-0311551P.
 PR 08-NOV-2001; 2002US-0363649P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Thornton M, Au-Young JK, Azimzai Y, Bandman O, Barroso I;
 PI Baughn MR, Becha SD, Borowsky ML, Ding L, Duggan BM, Elliott VS;
 PI Emerling BM, Forsythe IJ, Gandhi AR, Gietzen KJ, Gorvad AE;
 PI Griffin JA, Gururajan R, Hafalia AJA, Ring HZ, Ison CH, Jones KA;
 PI Lal PG, Lee EA, Lee S, Li JX, Lu DAM, Marquis JP, Lehr-Mason PM;
 PI Chawla NK, Arvizu CS, Sanjanwala B, Sornasse T, Swarnakar A;
 PI Tang YT, Thangavelu K, Tran B, Tran UK, Warren BA, Xu Y, Yao MG;
 PI Yue H, Yue H, Zebardjian Y, Chang H;
 XX
 DR WPI; 2003-533003/50.
 DR N-PSDB; ABT43746.

XX New Molecules for Disease Detection and Treatment polypeptides, useful
 PT for preparing a composition for diagnosing or treating e.g.
 PT cardiovascular or neurological disorders.
 PS
 XX Claim 1; Page 239; 289pp; English.
 CC This invention relates to novel polypeptides associated with Molecules
 CC for Disease Detection and Treatment (MDDT) and the cDNA sequences which
 CC encode them. MDDT-agonists, MDDT-agonists or gene therapy may produce
 CC cardiant, cytostatic and neuroprotective activities. The polypeptides and
 CC cDNA sequences of the invention may be useful for the preparation of a
 CC composition for the diagnosis or treatment of a disease which is
 CC associated with decreased expression or over-expression of functional
 CC MDDT proteins, for example cardiovascular, neurological, cell
 CC proliferative, autoimmune/inflammatory or developmental disorders. They
 CC may also be useful in the treatment and diagnosis of cancer, diseases
 CC treated with steroids and disorders caused by the metabolic response to
 CC treatment with steroids. The present sequence is the amino acid sequence
 CC of a human Molecule for Disease Detection and Treatment (MDDT) of the
 CC invention
 XX
 SQ Sequence 305 AA;
 Query Match 60.0%; Score 39; DB 6; Length 305;
 Best Local Similarity 50.0%; Pred. No. 2.9e+02;
 Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 1 NSWDSGGTHP 10
 :|||
 Db 254 HQWDEQGNHP 263
 :|||
 RESULT 21
 AAB71910
 ID AAB71910 standard; protein; 311 AA.
 XX
 AC AAB71910;
 XX
 DT 09-MAY-2001 (first entry)
 XX
 DE Human ISOM-2.
 XX
 XX Human; ISOM; isomerase; antiallergic; dermatological;
 KW antidiabetic; immunosuppressive; cytostatic; antiinflammatory;
 KW osteopathic; antipsoriatic; antirheumatic; antiarthritic; anti-HIV;
 KW antiarteriosclerotic; antithyroid; nephrotropic; neuroprotective;
 KW hepatotropic; gene therapy; immune disorder; autoimmune disease;
 KW diabetes mellitus; cancer; cell proliferative disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200112790-A2.
 XX
 PD 22-FEB-2001.
 XX
 PF 16-AUG-2000; 2000WO-US0252518.
 XX
 PR 17-AUG-1999; 99US-0149388P.
 XX
 XX (INCY-) INCYTE GENOMICS INC.
 PA
 XX Bandman O, Lu DAM, Yue H, Tran B, Hillman JL, Baughn MR, Lal P;
 PI Tang YT;
 PI WPI; 2001-202861/20.
 DR N-PSDB; AAF60311.
 DR
 XX Novel human isomerases useful for diagnosing, preventing and treating
 PT immune disorders, inflammatory disorders and cell proliferative disorders
 PT including cancer.
 XX
 XX Claim 1; Page 88-89; 102pp; English.

XX
 CC The present sequence is an isolated human isomerase (ISOM) polypeptide.
 CC It is useful for treating a disease or condition associated with
 CC decreased or increased expression of functional human isomerase. It is
 CC useful for treating immune disorders such as inflammation, acquired
 CC immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory
 CC distress syndrome (ARDS), allergies, anaemia, asthma, atherosclerosis,
 CC Crohn's disease, atopic dermatitis, glomerulonephritis, Grave's disease,
 CC Hashimoto's thyroiditis, multiple sclerosis, osteoporosis, rheumatoid
 CC arthritis, psoriasis, myocardial or pericardial inflammation, and
 CC autoimmune diseases, diabetes mellitus, cancers and cell proliferative
 CC disorders such as actinic keratosis, prostatitis, cirrhosis, and
 CC myelofibrosis
 XX
 SQ Sequence 311 AA;
 Query Match 60.0%; Score 39; DB 4; Length 311;
 Best Local Similarity 66.7%; Pred. No. 3e+02;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 NSWDSGGTH 9
 :|||
 Db 32 NSWDLPGSH 40
 :|||
 RESULT 22
 AAU28081
 ID AAU28081 standard; protein; 324 AA.
 XX
 AC AAU28081;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Novel human secretory protein, Seq ID No 250.
 XX
 KW Human; secreted protein; arthritis; Crohn's disease; sepsis; shock;
 KW ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;
 KW transgenic animal; Alzheimer's disease; Parkinson's disease; burn;
 KW amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;
 KW ulcer; osteoporosis; bone degenerative disorder; periodontal disease;
 KW gut protection; lung; liver fibrosis; immune deficiency; infection;
 KW severe combined immunodeficiency; SCID; autoimmune disorder; allergy;
 KW multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;
 KW fertility; analgesic; pain; antigen.
 XX
 OS Homo sapiens.
 XX
 PN WO200166689-A2.
 XX
 PD 13-SEP-2001.
 XX
 PF 05-MAR-2001; 2001WO-US004942.
 XX
 PR 07-MAR-2000; 2000US-00519705.
 PR 19-MAY-2000; 2000US-00574454.
 PR 17-JUN-2000; 2000US-00596193.
 PR 14-JUL-2000; 2000US-00616847.
 PR 19-SEP-2000; 2000US-00665363.
 PR 20-OCT-2000; 2000US-00693267.
 XX
 XX (HYSE-) HYSEQ INC.
 PA
 XX Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;
 PI Zhao QA, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AJ, Wang J;
 PI WPI; 2001-589934/56.
 DR N-PSDB; AAS44981.
 DR
 XX Novel polypeptides and nucleic acids obtained from cDNA libraries
 PT prepared from various human tissues, for diagnosis and treatment of
 PT cancer, neurological, inflammatory, and autoimmune disorders.
 XX
 XX Example 4; SEQ ID NO 250; 107pp; English.

XX The invention relates to novel isolated human secreted polypeptides (I)
CC and polynucleotides (II). (I) and (II) are useful for treating
CC inflammatory conditions such as shock, sepsis, nephritis, Crohn's disease,
CC ischaemia-reperfusion injury, shock, sepsis, immune responses, and is
CC involved in increasing hematopoiesis, stem cell survival, bone growth
CC and remodeling. (I), (II) and modulators of (II) are useful for
CC prophylaxis or treatment of one or more cancers. (II) is also useful for
CC creating transgenic animals useful for studying the in vivo activities of
CC the polypeptide as well as for studying modulators of the polypeptides.
CC (I) induces the proliferation of neural cells and regeneration of nerve
CC and brain tissue and is useful for the treatment of central and
CC peripheral nervous system diseases and neuropathies, such as Alzheimer's,
CC Parkinson's disease, Huntington's disease, and amyotrophic lateral
CC sclerosis. In addition, (I) is involved in chemotactic or chemokinetic
CC activity, regulation of hematopoiesis and is useful for treating myeloid
CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia
CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve
CC tissue growth, and in tissue repair, healing of burns, incisions, ulcers,
CC for treating osteoporosis, osteoarthritis, bone degenerative disorders,
CC or periodontal disease. Furthermore, (I) is also useful for gut
CC protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues, various immune deficiencies and
CC disorders including severe combined immunodeficiency (SCID), bacterial or
CC fungal infections, autoimmune disorders e.g. multiple sclerosis,
CC rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic
CC reactions and conditions, such as asthma or other respiratory problems.
CC In addition, (I) affects biorhythms or circadian cycles of rhythms,
CC fertility, metabolism, catabolism, anabolism, storage or elimination of
CC dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides
CC analgesic effects or other pain reducing effects, immunoglobulin like
CC activity and can act as an antigen in a vaccine composition to raise an
CC immune response. AAU28020-AAU28395 represent novel human secreted protein
CC amino acid sequences of the invention
XX
XX Sequence 324 AA;
SQ

Query Match 60.0%; Score 39; DB 4; Length 324;
Best Local Similarity 77.8%; Pred. No. 3.1e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 SWDSSGTHP 10
Db 162 SWDFSGTTP 170
||| |||

RESULT 23
ABU03132
XX ABU03132 standard; protein; 683 AA.
XX AC ABU03132;
XX DT 21-JAN-2003 (first entry)
XX DE Alpha amylase polypeptide #93.
XX KW Alpha amylase; enzyme; starch linkage hydrolysis; starch liquefaction;
KW starch breakdown catalysis; textile desizing; lignocellulosic fibre;
KW enzymatic de-linking; recycled paper; high-maltose syrup; dough;
KW high glucose syrup; corn-wet milling process; detergent; baking process;
KW beverage; oil field; fuel ethanol; brewing process; staling;
KW starch modification.
XX OS Unidentified.
XX FN WO200268589-A2.
XX PD 06-SEP-2002.
XX PP 21-FEB-2002; 2002WO-US005068.
XX PR 21-FEB-2001; 2001US-0270495P.
XX PR 21-FEB-2001; 2001US-0270496P.

PR 14-MAY-2001; 2001US-0291122P.
XX (DIVE-) DIVERGA CORP.
XX PI Callen W, Richardson T, Frey G;
XX DR WPI; 2003-018656/01.
XX DR N-PSDB; ABX08503.
XX Novel purified polypeptide with alpha-amylase activity, useful e.g. for
PT liquefying starch, for textile desizing, for treating lignocellulosic
PT fibers, and for producing high-maltose or high-glucose syrup.
XX Claim 30; Fig 16; 301pp; English.
XX The invention relates to a purified polypeptide with alpha-amylase
CC activity and the polynucleotide encoding it. The polypeptide is useful
CC for hydrolysing starch linkages, for catalysing the breakdown of a
CC starch, for modifying small molecules, for liquefying starch, for washing
CC an object, for textile desizing, for treating lignocellulosic fibers, for
CC improving fibre properties, for enzymatic de-linking of recycled paper
CC pulp, for producing a high-maltose or high-glucose syrup or a mixed
CC syrup, and for increasing the flow of production fluids from a
CC subterranean formation by removing a viscous, starch-containing, damaging
CC fluid formed during production operations and found within the
CC subterranean formation which surrounds a completed well bore. The
CC polypeptide is also useful for preparing a dough or a baked product
CC prepared from the dough and in corn-wet milling processes, detergents,
CC baking processes, beverages, oil fields (fuel ethanol), brewing processes
CC and starch modification in the paper and pulp industry, for removing
CC starch containing stains from a material and for reducing staling of
CC bakery products. Sequences ABU03040-ABU03144 represent alpha amylase
CC polypeptides of the invention
XX
XX Sequence 683 AA;
SQ

Query Match 60.0%; Score 39; DB 6; Length 683;
Best Local Similarity 66.7%; Pred. No. 7e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 NSWSSGTH 9
Db 364 NNWDYLGTH 372
||| |||

RESULT 24
AAU92973
XX AAU92973 standard; protein; 1336 AA.
XX AC AAU92973;
XX DT 02-JUL-2002 (first entry)
XX DE Arabidopsis transcription factor #11.
XX KW Agriculture; metabolic chemical; environmental stress; drought;
KW microbial disease resistance; herbicide resistance; seed yield;
KW fruit yield; growth rate; leaf senescence; flower senescence. Plant;
KW transcription factor; transgenic.
XX OS Arabidopsis thaliana.
XX FN WO200215675-A1.
XX PD 28-FEB-2002.
XX PF 22-AUG-2001; 2001WO-US026189.
XX PP 22-AUG-2000; 2000US-0227439P.
XX PR 16-NOV-2000; 2000US-00713994.
XX PR 18-APR-2001; 2001US-00837944.
XX PA (MEND-) MENDEL BIOTECHNOLOGY INC.

PA (PILG/) PILGRIM M.
PA (CREE/) CREELMAN R.
PA (DUBE/) DUBELL A. J.
PA (HEAR/) HEARD J.
PA (JIAN/) JIANG C.
PA (KEDD/) KEDDIE J.
PA (ADAM/) ADAM L.
PA (RATC/) RATCLIFFE O.
PA (REUB/) REUBER J. L.
PA (RIEC/) RIECHMANN J. L.
PA (YUGG/) YU G.
PA (PINE/) PINEDA O.
XX
PI Pilgrim M, Creelman R, Dubell AJ, Heard J, Jiang C, Keddle J;
PI Adam L, Ratcliffe O, Reuber JL, Riechmann JL, Yu G, Pineda O;
XX
DR WPI; 2002-292022/33.
DR N-PSDB; ABK65159.
XX
XX An isolated or recombinant polynucleotide used to produce a transgenic plant.
PT
PS Claim 40; Page 118-123; 941pp; English.
XX
CC The invention relates to 1 of 232 isolated or recombinant polynucleotides encoding an Arabidopsis thaliana transcription factor, their variants, complements, fragments, or related polynucleotide with 31% to 95% sequence identity, where the plant possesses an altered trait as compared to a wild-type or reference plant, or the plant exhibits an altered phenotype as compared to a wild-type or reference plant, or the plant exhibits ectopic expression or altered expression of one or more genes associated with a plant trait as compared to a wild plant. Also included are a transgenic plant comprising the polynucleotides, a computer readable medium having stored sequence information, and identifying a homologue sequence from a database comprising a plurality of known plant sequences comprising inputting sequence information selected from one of 464 fully defined sequences given in the specification. The isolated or recombinant polynucleotide is used for producing a plant having a modified trait, the method comprising selecting a polynucleotide that encodes a polypeptide or an antisense nucleic acid, inserting the polynucleotide or antisense nucleic acid into an expression vector, introducing the vector into a plant or a cell of a plant to overexpress the polypeptide or antisense nucleic acid, thereby producing a modified plant, and selecting for a modified trait (e.g. increased production of agriculturally useful proteins or metabolic chemicals, pest tolerance, environmental stress response (e.g. drought), microbial disease resistance, herbicide resistance, seed and fruit yield, growth rate, leaf and flower senescence and many other traits listed in the specification).
CC The present sequence is one of the 232 proteins which are A. thaliana transcription factors
XX
SQ Sequence 1336 AA;
XX
Query Match 60.0%; Score 39; DB 5; Length 1336;
Best Local Similarity 75.0%; Pred. No. 1.5e+03;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 NGWDSSTG 8
Db 798 NSWDASGS 805

RESULT 25
ADA15573
ID ADA15573 standard; protein; 1336 AA.
XX
AC ADA15573;
XX
DT 06-NOV-2003 (first entry)
XX
DE A. thaliana protein, SEQ ID NO:128.
XX
KW Transgenic plant; plant; transcription factor; trait; root;

KW plant stability.
XX Arabidopsis thaliana.
OS
XX US2003061637-A1.
PN
XX 27-MAR-2003.
PD
XX 21-OCT-2002; 2002US-00278173.
PF
XX 23-MAR-1999; 99US-0125814P.
PR
XX 22-MAR-2000; 2000US-00533393.
PR
XX (JIAN/) JIANG C.
PA (BROU/) BROUN P.
PA (RIEC/) RIECHMANN J L.
PA (PINE/) PINEDA O.
PA (ZHAN/) ZHANG J.
PA (YUGG/) YU G.
PA (PILG/) PILGRIM M.
PA (KEDD/) KEDDIE J.
PA (HEAR/) HEARD J.
PA (REUB/) REUBER L.
PA (RATC/) RATCLIFFE O.
PA (ADAM/) ADAM L.
PA (SAMA/) SAWAHA R.
XX
XX Jiang C, Broun P, Riechmann JL, Pineda O, Zhang J, Yu G;
PI Pilgrim M, Keddle J, Heard J, Reuber L, Ratcliffe O, Adam L;
PI Samaha R;
XX
XX WPI; 2003-555503/52.
DR N-PSDB; ADA15572.
DR
XX New transgenic plant comprising a recombinant polynucleotide, useful for altering a plant's trait for increasing plant stability.
PT
PT
XX Claim 1; SEQ ID NO 128; 159pp; English.
PS
XX The invention discloses a transgenic plant having an A. thaliana recombinant polynucleotide (a transcription factor) comprising a sequence encoding a polypeptide with at least 6 consecutive amino acids of one of the amino acid sequences given in the specification. The recombinant polynucleotide alters a trait of the transgenic plant's roots when compared to the same trait of the roots of another plant lacking the recombinant polynucleotide. Also claimed are methods for altering the expression levels of at least one gene of a plant, altering a trait associated with a plant's roots and altering a plant's trait. The recombinant polynucleotide further comprises a promoter operably linked to the nucleotide sequence. The promoter is constitutive or inducible or root-active. The method for altering a trait associated with roots comprises transforming a plant with the recombinant polynucleotide, selecting the transformed plants and identifying a transformed plant with roots having an altered trait. The method for altering the expression levels of at least one gene of a plant comprises transforming a plant with the recombinant polynucleotide and selecting the transformed plant. The method for altering a plant's trait comprises providing a database sequence comparing the database sequence with the polypeptide, selecting the database that meets selected sequence criteria and transforming the selected database sequence in the plant. The methods also comprise providing a test polynucleotide, hybridising the test polynucleotide with the recombinant polynucleotide at low stringency and transforming the hybridising test polynucleotide in a plant to alter a trait of the plant. The transgenic plant is useful for altering a plant's trait for increasing plant stability. The sequence presented is a protein of the invention.
XX
SQ Sequence 1336 AA;
XX
Query Match 60.0%; Score 39; DB 6; Length 1336;
Best Local Similarity 75.0%; Pred. No. 1.5e+03;
XX

Matches	6;	Conservative	2;	Mismatches	0;	Indels	0;	Gaps	0;
QY	1	NSWDSSGT 8							
Db	798	NSWDASGS 805							

Search completed: September 27, 2004, 12:13:28
Job time : 127 secs

This Page Blank (uspio)

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	40	61.5	448	4	US-09-252-991A-24066	Sequence 24066, A
2	38	58.5	92	2	US-08-447-464-6	Sequence 6, Appli
3	38	58.5	92	2	US-08-716-679-6	Sequence 6, Appli
4	38	58.5	461	1	US-08-672-571A-3	Sequence 3, Appli
5	38	58.5	490	1	US-08-672-571A-1	Sequence 1, Appli
6	38	58.5	1452	2	US-08-652-971-A	Sequence 4, Appli
7	38	58.5	1452	2	US-08-449-644-8	Sequence 8, Appli
8	38	58.5	1452	2	US-08-087-244A-8	Sequence 8, Appli
9	38	58.5	1452	2	US-08-991-258A-4	Sequence 4, Appli
10	38	58.5	1452	2	US-08-769-399-4	Sequence 4, Appli
11	38	58.5	1452	3	US-08-991-953A-4	Sequence 4, Appli
12	37.5	57.7	749	4	US-09-562-737-97	Sequence 97, Appli
13	37	56.9	11	2	US-08-350-260A-349	Sequence 349, App
14	37	56.9	11	2	US-08-350-260A-430	Sequence 430, App
15	37	56.9	11	2	US-09-104-337A-349	Sequence 349, App
16	37	56.9	11	4	US-09-104-337A-430	Sequence 430, App
17	37	56.9	23	4	US-09-205-258-1061	Sequence 1061, Ap
18	37	56.9	32	3	US-09-167-681-34	Sequence 34, Appli
19	37	56.9	32	3	US-09-167-681-42	Sequence 42, Appli
20	37	56.9	32	3	US-09-167-681-50	Sequence 50, Appli
21	37	56.9	32	2	US-08-273-146-67	Sequence 67, Appli
22	37	56.9	95	4	US-09-621-976-7103	Sequence 7103, Ap
23	37	56.9	97	2	US-08-665-202-35	Sequence 35, Appli
24	37	56.9	97	4	US-09-315-574-35	Sequence 35, Appli
25	37	56.9	103	4	US-09-673-395A-348	Sequence 348, App
26	37	56.9	109	2	US-08-652-816A-16	Sequence 16, Appli
27	37	56.9	118	4	US-09-205-258-1060	Sequence 1060, Ap

101 34 52.3 166 3 US-08-384-106A-22 Sequence 22, Appl
102 34 52.3 166 3 US-09-240-906-2 Sequence 2, Appl
103 34 52.3 166 3 US-09-240-906-4 Sequence 4, Appl
104 34 52.3 166 3 US-09-483-597-2 Sequence 2, Appl
105 34 52.3 166 4 US-09-483-597-4 Sequence 4, Appl
106 34 52.3 166 5 PCT-US96-01643-2 Sequence 2, Appl
107 34 52.3 166 5 PCT-US96-01643-9 Sequence 9, Appl
108 34 52.3 166 5 PCT-US96-05252-2 Sequence 2, Appl
109 34 52.3 166 5 PCT-US96-05252-4 Sequence 4, Appl
110 34 52.3 204 4 US-09-489-039A-7286 Sequence 7286, Ap
111 34 52.3 214 3 US-08-861-774E-32 Sequence 32, Appl
112 34 52.3 234 4 US-09-252-991A-21695 Sequence 21695, A
113 34 52.3 247 4 US-09-107-532A-3856 Sequence 3856, Ap
114 34 52.3 264 3 US-09-053-197A-8 Sequence 8, Appl
115 34 52.3 264 4 US-09-085-761A-8 Sequence 8, Appl
116 34 52.3 283 4 US-09-798-051-7 Sequence 7, Appl
117 34 52.3 286 4 US-09-489-039A-8008 Sequence 8008, Ap
118 34 52.3 298 4 US-09-543-681A-8224 Sequence 8224, Ap
119 34 52.3 312 4 US-09-107-532A-6637 Sequence 6637, Ap
120 34 52.3 324 4 US-09-134-000C-6086 Sequence 6086, Ap
121 34 52.3 325 3 US-09-124-758-6 Sequence 6, Appl
122 34 52.3 325 4 US-09-709-677-6 Sequence 6, Appl
123 34 52.3 338 4 US-09-252-991A-27370 Sequence 27370, A
124 34 52.3 379 4 US-09-252-991A-28086 Sequence 28086, A
125 34 52.3 382 4 US-09-328-352-4850 Sequence 4850, Ap
126 34 52.3 397 4 US-09-252-991A-22719 Sequence 22719, A
127 34 52.3 400 3 US-08-725-459B-19 Sequence 19, Appl
128 34 52.3 426 4 US-09-134-000C-4710 Sequence 4710, Ap
129 34 52.3 496 4 US-09-252-991A-26506 Sequence 26506, A
130 34 52.3 535 3 US-08-725-459B-18 Sequence 18, Appl
131 34 52.3 609 3 US-08-969-683A-67 Sequence 67, Appl
132 34 52.3 619 4 US-09-489-039A-10209 Sequence 10209, A
133 34 52.3 630 3 US-08-725-459B-17 Sequence 17, Appl
134 34 52.3 680 3 US-08-725-459B-3 Sequence 3, Appl
135 34 52.3 683 4 US-09-213-293D-1 Sequence 1, Appl
136 34 52.3 730 3 US-08-725-459B-2 Sequence 2, Appl
137 34 52.3 733 3 US-08-725-459B-21 Sequence 21, Appl
138 34 52.3 733 3 US-08-725-459B-22 Sequence 22, Appl
139 34 52.3 733 3 US-08-725-459B-23 Sequence 23, Appl
140 34 52.3 733 3 US-08-725-459B-28 Sequence 28, Appl
141 34 52.3 733 3 US-08-725-459B-29 Sequence 29, Appl
142 34 52.3 733 3 US-08-725-459B-30 Sequence 30, Appl
143 34 52.3 748 3 US-08-725-459B-24 Sequence 24, Appl
144 34 52.3 748 3 US-08-725-459B-27 Sequence 27, Appl
145 34 52.3 753 3 US-08-725-459B-25 Sequence 25, Appl
146 34 52.3 753 3 US-08-725-459B-26 Sequence 26, Appl
147 34 52.3 769 3 US-08-725-459B-35 Sequence 35, Appl
148 34 52.3 769 3 US-08-725-459B-36 Sequence 36, Appl
149 34 52.3 769 3 US-08-725-459B-37 Sequence 37, Appl
150 34 52.3 769 3 US-08-725-459B-38 Sequence 38, Appl

ALIGNMENTS

RESULT 1
US-09-252-991A-24066
; Sequence 24066, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24066
; LENGTH: 448

; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24066
Query Match 61.5%; Score 40; DB 4; Length 448;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 DSSGTHP 10
Db 252 DSSGTHP 258
||| |||
RESULT 2
US-08-447-464-6
; Sequence 6, Application US/08447464
; Patent No. 5840842
; GENERAL INFORMATION:
; APPLICANT: Schlensing, Joseph
; APPLICANT: Yan, Hai
; TITLE OF INVENTION: NOVEL RECEPTOR-TYPE PROTEIN
; TITLE OF INVENTION: PHOSPHOTYROSINE PHOSPHATASE-SIGMA
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/447,464
; FILING DATE: 24-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/130,570
; FILING DATE: 01-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7683-043
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-969-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 92 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-447-464-6
Query Match 58.5%; Score 38; DB 2; Length 92;
Best Local Similarity 60.0%; Pred. No. 26;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 2 SWDSSGTHPV 11
Db 19 TWDSGNTEPV 28
||| |||
RESULT 3
US-08-716-679-6
; Sequence 6, Application US/08716679
; Patent No. 5846800
; GENERAL INFORMATION:

```

; APPLICANT: Schlössinger, Joseph
; APPLICANT: Yan, Hai
; TITLE OF INVENTION: NOVEL RECEPTOR-TYPE PROTEIN
; TITLE OF INVENTION: PHOSPHOTYROSINE PHOSPHATASE-SIGMA
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/716,679
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/130,570
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Mierock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7683-043
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 92 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-716-679-6

Query Match 58.5%; Score 38; DB 2; Length 92;
Best Local Similarity 60.0%; Pred. No. 26;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 SWDSSGTHPV 11
DB 19 TWDSNGTEPV 28

RESULT 4
US-08-672-571A-3
; Sequence 3, Application US/08672571A
; Patent No. 5795765
; GENERAL INFORMATION:
; APPLICANT: IZU, Hiroyuki
; APPLICANT: KURUME, Yoko
; APPLICANT: IZUMI, Yoshiya
; APPLICANT: SANO, Mutsumi
; APPLICANT: KATO, Ikunoshin
; APPLICANT: ITO, Makoto
; TITLE OF INVENTION: Gene Encoding Endoglycoceramidase
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/672,571A
; FILING DATE: 28 JUNE 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: WEINER, Marc S.
; REGISTRATION NUMBER: 32,181
; REFERENCE/DOCKET NUMBER: 1422-0264P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 490 amino acids
; TYPE: amino acid
; MOLECULE TYPE: IBM PC compatible

```

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/672,571A
; FILING DATE: 28 JUNE 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: WEINER, Marc S.
; REGISTRATION NUMBER: 32,181
; REFERENCE/DOCKET NUMBER: 1422-0264P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 461 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-672-571A-3

Query Match 58.5%; Score 38; DB 1; Length 461;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 NSWSSSGTHP 10
DB 164 NFWNTTGKHP 173

RESULT 5
US-08-672-571A-1
; Sequence 1, Application US/08672571A
; Patent No. 5795765
; GENERAL INFORMATION:
; APPLICANT: IZU, Hiroyuki
; APPLICANT: KURUME, Yoko
; APPLICANT: IZUMI, Yoshiya
; APPLICANT: SANO, Mutsumi
; APPLICANT: KATO, Ikunoshin
; APPLICANT: ITO, Makoto
; TITLE OF INVENTION: Gene Encoding Endoglycoceramidase
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/672,571A
; FILING DATE: 28 JUNE 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: WEINER, Marc S.
; REGISTRATION NUMBER: 32,181
; REFERENCE/DOCKET NUMBER: 1422-0264P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 490 amino acids
; TYPE: amino acid

```

✓

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/087,244A
FILING DATE: 01-JUL-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7683-042
TELEPHONE: 212-790-9090
TELEFAX: 212-869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1452 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-087-244A-8

Query Match 58.5%; Score 38; DB 2; Length 1452;
Best Local Similarity 55.6%; Pred. No. 4.8e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 SWDSSGTHP 10
Db 434 SWDTNSHP 442

RESULT 9
US-08-991-258A-4
Sequence 4, Application US/08991258A
Patent No. 5928887
GENERAL INFORMATION:
APPLICANT: Cheng, Jill
APPLICANT: Lasky, Laurence A.
TITLE OF INVENTION: A NOVEL KAPPA/MU-LIKE PROTEIN TYROSINE
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSES: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT, LLP
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/991,258A
FILING DATE: 17-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/652,971
FILING DATE: 24-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Dreger, Walter H.
REGISTRATION NUMBER: 24,190
REFERENCE/DOCKET NUMBER: A-63478-3/WH/MTK
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1452 amino acids

TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-991-258A-4

Query Match 58.5%; Score 38; DB 2; Length 1452;
Best Local Similarity 55.6%; Pred. No. 4.8e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 SWDSSGTHP 10
Db 434 SWDTNSHP 442

RESULT 10
US-08-769-399-4
Sequence 4, Application US/08769399
Patent No. 5976852
GENERAL INFORMATION:
APPLICANT: Cheng, Jill
APPLICANT: Lasky, Laurence A.
TITLE OF INVENTION: A NOVEL KAPPA/MU-LIKE PROTEIN TYROSINE
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSES: Genentech, Inc.
STREET: 460 Point San Bruno Blvd.
CITY: South San Francisco
STATE: California
COUNTRY: United States
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/769,399
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Dreger, Ginger R.
REGISTRATION NUMBER: 33,055
REFERENCE/DOCKET NUMBER: P1033
TELEPHONE: (415) 225-3216
TELEFAX: (415) 952-9881
TELEX: 910 371-7168
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1452 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-769-399-4

Query Match 58.5%; Score 38; DB 2; Length 1452;
Best Local Similarity 55.6%; Pred. No. 4.8e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 SWDSSGTHP 10
Db 434 SWDTNSHP 442

RESULT 11
US-08-991-953A-4
Sequence 4, Application US/08991953A
Patent No. 6083748
GENERAL INFORMATION:
APPLICANT: Cheng, Jill

APPLICANT: Laskey, Laurence A.
TITLE OF INVENTION: A NOVEL KAPPA/MU-LIKE PROTEIN TYROSINE
NUMBER OF INVENTION: PHOSPHATASE, PTP LAMBDA
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOHACH, TEST, ALBRITTON & HERBERT, LLP
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/991,953A
FILING DATE: 16-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/652,971
FILING DATE: 24-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Dreger, Walter H.
REGISTRATION NUMBER: 24,190
REFERENCE/DOCKET NUMBER: A-63478-3/WH/MTK
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1452 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-991-953A-4

Query Match 58.5%; Score 38; DB 3; Length 1452;
Best Local Similarity 55.6%; Pred. No. 4.8e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 SWDSSGTHP 10
|||: |||
DB 434 SWDTNSHP 442

RESULT 12
US-09-562-737-97
Sequence 97, Application US/09562737
Patent No. 6428967
GENERAL INFORMATION:
APPLICANT: Herz, Joachim
APPLICANT: Gotthardt, Michael
TITLE OF INVENTION: LDL Receptor Signaling Pathways
FILE REFERENCE: UTSW0708
CURRENT APPLICATION NUMBER: US/09/562,737
CURRENT FILING DATE: 2000-05-01
NUMBER OF SEQ ID NOS: 132
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 97
LENGTH: 749
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
OTHER INFORMATION: Sequence
US-09-562-737-97

Query Match 57.7%; Score 37.5; DB 4; Length 749;
Best Local Similarity 77.8%; Pred. No. 2.9e+02;

Matches 7; Conservative 1; Mismatches 0; Indels 1; Gaps 1;
QY 3 W-DSSGTHP 10
|||: |||
DB 141 WTDSAGTHP 149

RESULT 13
US-08-350-260A-349
Sequence 349, Application US/08350260A
Patent No. 5962255
GENERAL INFORMATION:
APPLICANT: Winter, Gregory Paul
APPLICANT: Griffiths, Andrew David
APPLICANT: Williams, Samuel Cameron
APPLICANT: Waterhouse, Peter
APPLICANT: Nissim, Ahuva
APPLICANT: Johnson, Kevin Stuart
APPLICANT: Smith, Andrew John Hammond
TITLE OF INVENTION: Methods for producing members of specific
TITLE OF INVENTION: binding pairs
CORRESPONDENCE ADDRESS:
NUMBER OF SEQUENCES: 602
ADDRESSEE: David W. Clough
STREET: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606-6402
TELEX:
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/350,260A
FILING DATE: 05-DEC-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9110549.4
FILING DATE: 15-MAY-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9206318.9
FILING DATE: 24-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB91/01134
FILING DATE: 10-JUL-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB92/00883
FILING DATE: 15-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB93/00605
FILING DATE: 24-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/150,002
FILING DATE: 31-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/307,619
FILING DATE: 16-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 28111/32372
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
INFORMATION FOR SEQ ID NO: 349:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-350-260A-349

Query Match 56.9%; Score 37; DB 2; Length 11;
Best Local Similarity 72.7%; Pred. No. 4.1;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 NSWDSSGTHPV 11
||| ||| ||| |||
DB 1 NSRDSSGNHRV 11

RESULT 14
US-08-350-260A-430
; Sequence 430, Application US/08350260A
; Patent No. 5962255
; GENERAL INFORMATION:
; APPLICANT: Winter, Gregory Paul
; APPLICANT: Griffiths, Andrew David
; APPLICANT: Williams, Samuel Cameron
; APPLICANT: Waterhouse, Peter
; APPLICANT: Nissim, Ahuva
; APPLICANT: Johnson, Kevin Stuart
; APPLICANT: Smith, Andrew John Hammond
; TITLE OF INVENTION: Methods for producing members of specific
; TITLE OF INVENTION: binding pairs
; NUMBER OF SEQUENCES: 602
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David W. Clough
; STREET: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/350,260A
; FILING DATE: 05-DEC-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9110549.4
; FILING DATE: 15-MAY-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9206318.9
; FILING DATE: 24-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB91/01134
; FILING DATE: 10-JUL-1991
; PRIOR APPLICATION DATA: PCT/GB92/00883
; APPLICATION NUMBER:
; FILING DATE: 15-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB93/00605
; FILING DATE: 24-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/150,002
; FILING DATE: 31-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/307,619
; FILING DATE: 16-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Clough, David W
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 28111/32372
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; INFORMATION FOR SEQ ID NO: 430:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid

STRANDEDNESS: single
TOPOLOGY: linear
US-08-350-260A-430

Query Match 56.9%; Score 37; DB 2; Length 11;
Best Local Similarity 72.7%; Pred. No. 4.1;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 NSWDSSGTHPV 11
||| ||| ||| |||
DB 1 NSRDSSGNHRV 11

RESULT 15
US-09-104-337A-349
; Sequence 349, Application US/09104337A
; Patent No. 6492160
; GENERAL INFORMATION:
; APPLICANT: Winter, Gregory Paul
; APPLICANT: Griffiths, Andrew David
; APPLICANT: Williams, Samuel Cameron
; APPLICANT: Waterhouse, Peter
; APPLICANT: Nissim, Ahuva
; APPLICANT: Johnson, Kevin Stuart
; APPLICANT: Smith, Andrew John Hammond
; TITLE OF INVENTION: Methods for producing members of specific
; TITLE OF INVENTION: binding pairs
; NUMBER OF SEQUENCES: 600
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Audrey L. Bartnicki
; STREET: Marshall, Gerstein & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/104,337A
; FILING DATE: 25-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/350,260
; FILING DATE: 05-DEC-1994
; APPLICATION NUMBER: GB 9110549.4
; FILING DATE: 15-MAY-1991
; APPLICATION NUMBER: GB 9206318.9
; FILING DATE: 24-MAR-1992
; APPLICATION NUMBER: PCT/GB92/00883
; FILING DATE: 15-MAY-1992
; APPLICATION NUMBER: PCT/GB93/00605
; FILING DATE: 24-MAR-1993
; APPLICATION NUMBER: US 08/150,002
; FILING DATE: 31-MAR-1994
; APPLICATION NUMBER: US 08/307,619
; FILING DATE: 16-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bartnicki, Audrey L.
; REGISTRATION NUMBER: 40,499
; REFERENCE/DOCKET NUMBER: 28111/32372A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; INFORMATION FOR SEQ ID NO: 349:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 349:

US-09-104-337A-349


```
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,983
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1061
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-205-258-1061
```

```
Query Match 56.9%; Score 37; DB 4; Length 23;
Best Local Similarity 54.5%; Pred. No. 8.9;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 1 NSWDSGTHPV 11
Db 13 NATSSAGTHPL 23
```

```
RESULT 18
US-09-167-681-34
; Sequence 34, Application US/09167681A
; Patent No. 6265561
; GENERAL INFORMATION:
; APPLICANT: Weinshilboun, M.D., Richard M.
; APPLICANT: Raftogianis, Rebecca B.
; APPLICANT: Wood, Thomas C.
; APPLICANT: Otterness, Diane M.
; TITLE OF INVENTION: SULFOTRANSFERASE SEQUENCE VARIANTS
; FILE REFERENCE: 07039/118001
; CURRENT APPLICATION NUMBER: US/09/167,681A
; CURRENT FILING DATE: 1998-10-07
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 34
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-167-681-34
```

```
Query Match 56.9%; Score 37; DB 3; Length 32;
Best Local Similarity 66.7%; Pred. No. 13;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 3 WDSSGTHPV 11
Db 14 WELSRTHPV 22
```

```
RESULT 19
US-09-167-681-42
```

```
; Sequence 42, Application US/09167681A
; Patent No. 6265561
; GENERAL INFORMATION:
; APPLICANT: Weinshilboun, M.D., Richard M.
; APPLICANT: Raftogianis, Rebecca B.
; APPLICANT: Wood, Thomas C.
; APPLICANT: Otterness, Diane M.
; TITLE OF INVENTION: SULFOTRANSFERASE SEQUENCE VARIANTS
; FILE REFERENCE: 07039/118001
; CURRENT APPLICATION NUMBER: US/09/167,681A
; CURRENT FILING DATE: 1998-10-07
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 42
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-167-681-42
```

```
Query Match 56.9%; Score 37; DB 3; Length 32;
Best Local Similarity 66.7%; Pred. No. 13;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 3 WDSSGTHPV 11
Db 14 WELSRTHPV 22
```

```
RESULT 20
US-09-167-681-50
; Sequence 50, Application US/09167681A
; Patent No. 6265561
; GENERAL INFORMATION:
; APPLICANT: Weinshilboun, M.D., Richard M.
; APPLICANT: Raftogianis, Rebecca B.
; APPLICANT: Wood, Thomas C.
; APPLICANT: Otterness, Diane M.
; TITLE OF INVENTION: SULFOTRANSFERASE SEQUENCE VARIANTS
; FILE REFERENCE: 07039/118001
; CURRENT APPLICATION NUMBER: US/09/167,681A
; CURRENT FILING DATE: 1998-10-07
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 50
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-167-681-50
```

```
Query Match 56.9%; Score 37; DB 3; Length 32;
Best Local Similarity 66.7%; Pred. No. 13;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 3 WDSSGTHPV 11
Db 14 WELSRTHPV 22
```

```
RESULT 21
US-08-273-146-67
; Sequence 67, Application US/08273146
; Patent No. 5855885
; GENERAL INFORMATION:
; APPLICANT: Smith, Rodger
; APPLICANT: McCafferty, John
; APPLICANT: Chiswell, David
; APPLICANT: Fitzgerald, Michael J.
; APPLICANT: Fitzgerald, Kevin
; APPLICANT: Kenten, John H.
; APPLICANT: Martin, Mark T.
; APPLICANT: Titmas, Richard C.
; APPLICANT: Williams, Richard O.
; TITLE OF INVENTION: The Isolation and Production of
```

; TITLE OF INVENTION: Catalytic Antibodies using Phage Technology

; NUMBER OF SEQUENCES: 71
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IGEN, Inc.
; STREET: 1530 East Jefferson St.
; CITY: Rockville
; STATE: MD
; COUNTRY: USA
; ZIP: 20852
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/273,146
; FILING DATE: 14-JUL-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ryan, John W.
; REGISTRATION NUMBER: 33,771
; REFERENCE/DOCKET NUMBER: 09000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-984-8000
; TELEFAX: 301-230-0158
; INFORMATION FOR SEQ ID NO: 67:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 67 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-273-146-67

Query Match 56.9%; Score 37; DB 2; Length 67;
Best Local Similarity 72.7%; Pred. No. 27;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 NSWSSSGTHPV 11
||| |||||
DB 45 NSRDSGNHRV 55

RESULT 22

US-09-621-976-7103
; Sequence 7103, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 7103
; LENGTH: 95
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-621-976-7103

Query Match 56.9%; Score 37; DB 4; Length 95;
Best Local Similarity 66.7%; Pred. No. 40;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 NSWSSSGTH 9
:||| |||||
DB 19 SSWTSRGTH 27

RESULT 23

US-08-665-202-35

; Sequence 35, Application US/08665202
; Patent No. 5977322
; GENERAL INFORMATION:
; APPLICANT: Marks, James D.
; APPLICANT: Schier, Robert
; TITLE OF INVENTION: No. 5977322el High Affinity Human Antibodies to
; TITLE OF INVENTION: Tumor Antigens
; NUMBER OF SEQUENCES: 141
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/665,202
; FILING DATE: 13-JUN-1996
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/000,238
; FILING DATE: 14-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/000,250
; FILING DATE: 15-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Hunter, Tom
; REGISTRATION NUMBER: 38,498
; REFERENCE/DOCKET NUMBER: 02307E-061410
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 97 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-665-202-35

Query Match 56.9%; Score 37; DB 2; Length 97;
Best Local Similarity 72.7%; Pred. No. 41;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 NSWSSSGTHPV 11
||| |||||
DB 87 NSRDSGNHRV 97

RESULT 24

US-09-315-574-35
; Sequence 35, Application US/09315574
; Patent No. 6512097
; GENERAL INFORMATION:
; APPLICANT: Marks, James D.
; APPLICANT: Schier, Robert
; TITLE OF INVENTION: No. 6512097el High Affinity Human Antibodies to
; TITLE OF INVENTION: Tumor Antigens
; NUMBER OF SEQUENCES: 141
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Majestic, Parsons, Siebert & Haue P.C.
; STREET: Four Embarcadero Center, Suite 1100
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4106
; COMPUTER READABLE FORM:

Search completed: September 27, 2004, 12:17:14
Job time : 34 secs

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/315,574
FILING DATE: 20-MAY-99
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/000,238
FILING DATE: 14-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/000,250
FILING DATE: 15-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/665,202
FILING DATE: 13-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Hunter, Tom
REGISTRATION NUMBER: 38,498
REFERENCE/DOCKET NUMBER: 02307E-061411
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 97 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-315-574-35

Query Match 56.9%; Score 37; DB 4; Length 97;
Best Local Similarity 72.7%; Pred. No. 41;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 NSWSSSGTHPV 11
DB 87 NSRDSSGNHV 97

RESULT 25
US-09-673-395A-348
Sequence 348, Application US/09673395A
Patent No. 6620923
GENERAL INFORMATION:
APPLICANT: SPECHT, THOMAS
APPLICANT: HINZMANN, BERND
APPLICANT: SCHMITT, ARMIN
APPLICANT: PILARSKY, CHRISTIAN
APPLICANT: DAHL, EDGAR
APPLICANT: ROSENTHAL, ANDRE
TITLE OF INVENTION: HUMAN NUCLEIC ACID SEQUENCES FROM UTERUS TUMOR TISSUE
FILE REFERENCE: ALBRE-12
CURRENT APPLICATION NUMBER: US/09/673,395A
CURRENT FILING DATE: 2000-10-17
NUMBER OF SEQ ID NOS: 637
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 348
LENGTH: 103
TYPE: PRT
ORGANISM: Homo sapiens
US-09-673-395A-348

Query Match 56.9%; Score 37; DB 4; Length 103;
Best Local Similarity 55.6%; Pred. No. 43;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 WSSSGTHPV 11
DB 19 WNSLGPPI 27

This Page Blank (uspio)

Result No.	Score	Query		DB	ID	Description
		Match	%			
1	51	78.5	11	9	US-09-817-661-21	Sequence 21, Appl
2	59	75.4	110	12	US-10-251-0858-173	Sequence 173, Appl
3	49	75.4	110	16	US-10-737-252-173	Sequence 173, Appl
4	49	75.4	168	12	US-10-424-599-177279	Sequence 177279
5	46	70.8	249	10	US-09-880-748-1115	Sequence 1115, Ap
6	46	70.8	249	12	US-10-233-418-1115	Sequence 1115, Ap
7	44	67.7	247	10	US-09-880-748-1116	Sequence 1116, Ap
8	44	67.7	247	12	US-10-737-252-1116	Sequence 1116, Ap
9	43	66.2	245	16	US-10-779-461-10	Sequence 10, Appl
10	41	63.1	183	16	US-10-437-963-194069	Sequence 194069
11	41	63.1	246	10	US-09-880-748-1980	Sequence 1980, Ap
12	41	63.1	246	12	US-10-293-418-1980	Sequence 1980, Ap
13	40	61.5	13	10	US-09-563-222-25	Sequence 25, Appl
14	40	61.5	90	12	US-10-424-599-243244	Sequence 243244
15	40	61.5	112	14	US-10-340-536-18	Sequence 18, Appl

89 37 56.9 233 12 US-10-029-926B-27
90 37 56.9 233 15 US-10-032-037B-27
91 37 56.9 233 15 US-10-029-988B-27
92 37 56.9 233 15 US-10-032-423A-27
93 37 56.9 237 10 US-09-880-748-1941
94 37 56.9 237 10 US-09-880-748-2039
95 37 56.9 237 10 US-09-880-748-2112
96 37 56.9 237 12 US-10-293-418-1941
97 37 56.9 237 12 US-10-293-418-2039
98 37 56.9 237 12 US-10-293-418-2112
99 37 56.9 238 10 US-09-880-748-1907
100 37 56.9 238 10 US-09-880-748-1931
101 37 56.9 238 10 US-09-880-748-2024
102 37 56.9 238 12 US-10-293-418-1907
103 37 56.9 238 12 US-10-293-418-1931
104 37 56.9 238 12 US-10-293-418-2024
105 37 56.9 239 10 US-09-880-748-937
106 37 56.9 239 10 US-09-880-748-2015
107 37 56.9 239 10 US-09-880-748-2018
108 37 56.9 239 10 US-09-880-748-2035
109 37 56.9 239 10 US-09-880-748-2038
110 37 56.9 239 12 US-10-293-418-937
111 37 56.9 239 12 US-10-293-418-2015
112 37 56.9 239 12 US-10-293-418-2018
113 37 56.9 239 12 US-10-293-418-2035
114 37 56.9 239 12 US-10-293-418-2038
115 37 56.9 240 10 US-09-880-748-1391
116 37 56.9 240 10 US-09-880-748-1912
117 37 56.9 240 10 US-09-880-748-2013
118 37 56.9 240 12 US-10-293-418-1391
119 37 56.9 240 12 US-10-293-418-1912
120 37 56.9 240 12 US-10-293-418-2013
121 37 56.9 240 14 US-10-322-673-52
122 37 56.9 241 10 US-09-880-748-1303
123 37 56.9 241 10 US-09-880-748-1911
124 37 56.9 241 10 US-09-880-748-2031
125 37 56.9 241 10 US-09-880-748-2060
126 37 56.9 241 12 US-10-293-418-1303
127 37 56.9 241 12 US-10-293-418-1911
128 37 56.9 241 12 US-10-293-418-2031
129 37 56.9 241 12 US-10-293-418-2060
130 37 56.9 242 10 US-09-880-748-1424
131 37 56.9 242 10 US-09-880-748-1444
132 37 56.9 242 10 US-09-880-748-1678
133 37 56.9 242 10 US-09-880-748-1742
134 37 56.9 242 10 US-09-880-748-1762
135 37 56.9 242 10 US-09-880-748-2021
136 37 56.9 242 10 US-09-880-748-2046
137 37 56.9 242 10 US-09-880-748-2106
138 37 56.9 242 12 US-10-293-418-1424
139 37 56.9 242 12 US-10-293-418-1444
140 37 56.9 242 12 US-10-293-418-1678
141 37 56.9 242 12 US-10-293-418-1742
142 37 56.9 242 12 US-10-293-418-1762
143 37 56.9 242 12 US-10-293-418-2021
144 37 56.9 242 12 US-10-293-418-2046
145 37 56.9 242 12 US-10-293-418-2106
146 37 56.9 243 10 US-09-880-748-1929
147 37 56.9 243 10 US-09-880-748-1940
148 37 56.9 243 10 US-09-880-748-1942
149 37 56.9 243 10 US-09-880-748-1944
150 37 56.9 243 10 US-09-880-748-2014

ALIGNMENTS

RESULT 1
US-09-817-661-21
; Sequence 21, Application US/09817661
; Patent No. US20020076692A1
; GENERAL INFORMATION:
; APPLICANT: Osbourn, Jane

; APPLICANT: Holet, Thor
; TITLE OF INVENTION: Improvements to ribosome display
; FILE REFERENCE: 84633
; CURRENT APPLICATION NUMBER: US/09/817,661
; CURRENT FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: US 60/193,802
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 21
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-817-661-21

Query Match 78.5%; Score 51; DB 9; Length 11;
Best Local Similarity 81.8%; Pred. No. 0.24;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NSWDSGTHPV 11
||| ||| |||
DB 1 NSWDSGNHV 11

RESULT 2

US-10-251-085B-173
; Sequence 173, Application US/10251085B
; Publication No. US20040072164A1
; GENERAL INFORMATION:
; APPLICANT: Bowdish, Katherine S.
; APPLICANT: Frederickson, Shana
; APPLICANT: Renshaw, Mark
; APPLICANT: Lin, Ying-Chi
; APPLICANT: Maruyama, Toshiaki
; TITLE OF INVENTION: ENGINEERED TEMPLATES AND THEIR USE IN SINGLE PRIMER AMPLIFICATION
; FILE REFERENCE: 1087-21
; CURRENT APPLICATION NUMBER: US/10/251,085B
; CURRENT FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 60/323,455
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 278
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 173
; LENGTH: 110
; TYPE: PRT
; ORGANISM: human
US-10-251-085B-173

Query Match 75.4%; Score 49; DB 12; Length 110;
Best Local Similarity 88.9%; Pred. No. 4.2;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 WDSSGTHPV 11
||| ||| |||
DB 92 WDSSGDHPV 100

RESULT 3

US-10-737-252-173
; Sequence 173, Application US/10737252
; Publication No. US20040175736A1
; GENERAL INFORMATION:
; APPLICANT: Bowdish, Katherine S.
; APPLICANT: Frederickson, Shana
; APPLICANT: Renshaw, Mark
; APPLICANT: Lin, Ying-Chi
; APPLICANT: Maruyama, Toshiaki
; TITLE OF INVENTION: ENGINEERED TEMPLATES AND THEIR USE IN SINGLE PRIMER AMPLIFICATION
; FILE REFERENCE: 1087-21 CIP
; CURRENT APPLICATION NUMBER: US/10/737,252
; CURRENT FILING DATE: 2003-12-15
; PRIOR APPLICATION NUMBER: US 10/251,085
; PRIOR FILING DATE: 2002-09-19

; PRIOR APPLICATION NUMBER: US 60/323,455
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 309
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 173
; LENGTH: 110
; TYPE: PRT
; ORGANISM: human
US-10-737-252-173

Query Match 75.4%; Score 49; DB 16; Length 110;
Best Local Similarity 88.9%; Pred. No. 4.2;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 WDSGTHPV 11
||| |||
DB 92 WDSGDHPV 100

RESULT 4.

US-10-424-599-177279
; Sequence 177279, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 177279
; LENGTH: 168
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_13109C.1.pep
US-10-424-599-177279

Query Match 75.4%; Score 49; DB 12; Length 168;
Best Local Similarity 70.0%; Pred. No. 6.2;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NSWDSGTHP 10
||| |||
DB 154 NSWESQSHP 163

RESULT 5

US-09-880-748-1115
; Sequence 1115, Application US/09880748
; Publication No. US20030059937A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PFS23
; CURRENT APPLICATION NUMBER: US/09/880,748
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 3239
; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 1115
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-748-1115

Query Match 70.8%; Score 46; DB 10; Length 249;
Best Local Similarity 72.7%; Pred. No. 26;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 NSWDSGTHPV 11
: ||| ||| |||
DB 228 HSWDSSGNHV 238

RESULT 6

US-10-293-418-1115
; Sequence 1115, Application US/10293418
; Publication No. US20030223996A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PFS23P2
; CURRENT APPLICATION NUMBER: US/10/293,418
; CURRENT FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: 60/331,469
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/340,817
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 09/880,748
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1115
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-418-1115

Query Match 70.8%; Score 46; DB 12; Length 249;
Best Local Similarity 72.7%; Pred. No. 26;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 NSWDSGTHPV 11
: ||| ||| |||
DB 228 HSWDSSGNHV 238

RESULT 7

US-09-880-748-1116
; Sequence 1116, Application US/09880748
; Publication No. US20030059937A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PFS23
; CURRENT APPLICATION NUMBER: US/09/880,748
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16

; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 3239
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1116
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-748-1116

Query Match 67.7%; Score 44; DB 10; Length 247;
Best Local Similarity 77.8%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NSWDSSGTH 9
:|||||||
Db 226 HSWDSSGNH 234

RESULT 8

US-10-293-418-1116
; Sequence 1116, Application US/10293418
; Publication No. US20030223956A1
; GENERAL INFORMATION:

; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523P2
; CURRENT APPLICATION NUMBER: US/10/293,418
; CURRENT FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: 60/331,469
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/340,817
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 09/880,748
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1116
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-418-1116

Query Match 67.7%; Score 44; DB 12; Length 247;
Best Local Similarity 77.8%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NSWDSSGTH 9
:|||||||
Db 226 HSWDSSGNH 234

RESULT 9

US-10-779-461-10
; Sequence 10, Application US/10779461
; Publication No. US20040166544A1
; GENERAL INFORMATION:

; APPLICANT: Morton, Philip A
; TITLE OF INVENTION: ANTIBODIES TO c-MET FOR THE TREATMENT OF CANCERS
; FILE REFERENCE: 00980/1
; CURRENT APPLICATION NUMBER: US/10/779,461
; CURRENT FILING DATE: 2004-02-13

; PRIOR APPLICATION NUMBER: 60/447,073
; PRIOR FILING DATE: 2003-02-13
; NUMBER OF SEQ ID NOS: 161
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 10
; LENGTH: 245
; TYPE: PRT
; ORGANISM: artificial
; FEATURE:
; OTHER INFORMATION: phage display generated human antibody
US-10-779-461-10

Query Match 66.2%; Score 43; DB 16; Length 245;
Best Local Similarity 77.8%; Pred. No. 77;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 WDSSGTHPV 11
:|||||||
Db 226 WDSSSDHPV 234

RESULT 10

US-10-437-963-194069
; Sequence 194069, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 194069
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_90147C.1.pap
US-10-437-963-194069

Query Match 63.1%; Score 41; DB 16; Length 183;
Best Local Similarity 70.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 NSWDSSGTHP 10
:|||||||
Db 49 NSWPESGTSF 58

RESULT 11

US-09-880-748-1980
; Sequence 1980, Application US/09880748
; Publication No. US20030059937A1
; GENERAL INFORMATION:

; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523
; CURRENT APPLICATION NUMBER: US/09/880,748
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16

; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 3239
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1980
; LENGTH: 246
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-748-1980

Query Match 63.1%; Score 41; DB 10; Length 246;
Best Local Similarity 81.8%; Pred. No. 1.6e+02;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NSWDSSGTHPV 11
||| |||||
Db 225 NSRDSGTHLV 235

RESULT 12
US-10-293-418-1980
; Sequence 1980, Application US/10293418
; Publication No. US20030223996A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind BLYS
; CURRENT APPLICATION NUMBER: US/10/293,418
; CURRENT FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: 60/331,469
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/340,817
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 09/880,748
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1980
; LENGTH: 246
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-418-1980

Query Match 63.1%; Score 41; DB 12; Length 246;
Best Local Similarity 81.8%; Pred. No. 1.6e+02;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NSWDSSGTHPV 11
||| |||||
Db 225 NSRDSGTHLV 235

RESULT 13
US-09-563-222-25
; Sequence 25, Application US/09563222
; Publication No. US20030079253A1
; GENERAL INFORMATION:
; APPLICANT: Hiatt, Andrew
; APPLICANT: Hein, Mich B.
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN ARRAYS IN
; TITLE OF INVENTION: EUKARYOTIC CELLS
; FILE REFERENCE: 310098.406

; CURRENT APPLICATION NUMBER: US/09/563,222
; CURRENT FILING DATE: 2000-05-02
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-563-222-25

Query Match 61.5%; Score 40; DB 10; Length 13;
Best Local Similarity 66.7%; Pred. No. 15;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 SWDSSGTHP 10
: |||||
Db 2 AWDSSSDHP 10

RESULT 14
US-10-424-599-243244
; Sequence 243244, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285884
; SEQ ID NO 243244
; LENGTH: 90
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_61678C.1.pap
US-10-424-599-243244

Query Match 61.5%; Score 40; DB 12; Length 90;
Best Local Similarity 66.7%; Pred. No. 89;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 NSWDSSGTH 9
||| |||||
Db 11 NSWDEGDDH 19

RESULT 15
US-10-340-536-18
; Sequence 18, Application US/10340536
; Publication No. US20030175212A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Rebecca
; APPLICANT: Born, Willi
; APPLICANT: Roark, Christina
; APPLICANT: Aydinug, M. Kemal
; TITLE OF INVENTION: Use of Soluble Gamma Delta T Cell Receptors for Regulating T Cell
; TITLE OF INVENTION: Function
; FILE REFERENCE: 2879-89
; CURRENT APPLICATION NUMBER: US/10/340,536
; CURRENT FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: 60/347,285
; PRIOR FILING DATE: 2002-01-10
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Mus musculus

US-10-340-536-18

Query Match 61.5%; Score 40; DB 14; Length 112;
Best Local Similarity 77.8%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 WDSSGTHPV 11
|||
DB 104 WDSSGFHKV 112

RESULT 16

US-10-340-536-16
; Sequence 16, Application US/10340536
; Publication No. US20030175212A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Rebecca
; APPLICANT: Born, Willi
; APPLICANT: Roark, Christina
; APPLICANT: Avdintug, M. Kemal
; TITLE OF INVENTION: Use of Soluble Gamma Delta T Cell Receptors for Regulating T Cell
; FILE OF INVENTION: Function
; FILE REFERENCE: 2879-89
; CURRENT APPLICATION NUMBER: US/10/340,536
; CURRENT FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: 60/347,285
; PRIOR FILING DATE: 2002-01-10
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 120
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-340-536-16

Query Match 61.5%; Score 40; DB 14; Length 120;
Best Local Similarity 77.8%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 WDSSGTHPV 11
|||
DB 112 WDSSGFHKV 120

RESULT 17

US-09-880-748-353
; Sequence 353, Application US/09880748
; Publication No. US20030059937A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523
; CURRENT APPLICATION NUMBER: US/09/880,748
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 3239
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 353
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-748-353

Query Match 61.5%; Score 40; DB 10; Length 249;

Best Local Similarity 88.9%; Pred. No. 2.3e+02;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NSWDSSGTH 9
|||
DB 228 NSRDSSGTH 236

RESULT 18

US-09-880-748-540
; Sequence 540, Application US/09880748
; Publication No. US20030059937A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523
; CURRENT APPLICATION NUMBER: US/09/880,748
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 3239
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 540
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-748-540

Query Match 61.5%; Score 40; DB 10; Length 249;
Best Local Similarity 88.9%; Pred. No. 2.3e+02;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NSWDSSGTH 9
|||
DB 228 NSRDSSGTH 236

RESULT 19

US-10-293-418-353
; Sequence 353, Application US/10293418
; Publication No. US20030223996A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523P2
; CURRENT APPLICATION NUMBER: US/10/293,418
; CURRENT FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: 60/331,469
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/340,817
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 09/880,748
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 353

; PRIOR FILING DATE: 2000-06-17
; PRIOR APPLICATION NUMBER: 09/574,454
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: 09/519,705
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 752
; SEQ ID NO 626
; LENGTH: 162
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-221-278-626

Query Match 60.0%; Score 39; DB 12; Length 162;
Best Local Similarity 77.8%; Pred. No. 2.2e+02; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 2;

QY 2 SWDSSGTHP 10
||| |||
Db 67 SWDFSGTTP 75

RESULT 24

US-10-291-172-626
; Sequence 626, Application US/10291172
; Publication No. US20030228584A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: No. US20030228584A1 Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-045
; CURRENT APPLICATION NUMBER: US/10/291,172
; CURRENT FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 09/693,267
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/665,363
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 09/616,847
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 09/596,193
; PRIOR FILING DATE: 2000-06-17
; PRIOR APPLICATION NUMBER: 09/574,454
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: 09/519,705
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 752
; SEQ ID NO 626
; LENGTH: 162
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-291-172-626

Query Match 60.0%; Score 39; DB 15; Length 162;
Best Local Similarity 77.8%; Pred. No. 2.2e+02; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 2;

QY 2 SWDSSGTHP 10
||| |||
Db 67 SWDFSGTTP 75

RESULT 25

US-10-437-963-164780
; Sequence 164780, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 164780
; LENGTH: 301
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_63647C.1.pap
US-10-437-963-164780

Query Match 60.0%; Score 39; DB 16; Length 301;
Best Local Similarity 75.0%; Pred. No. 3.9e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 2;

QY 3 WDSSGTHP 10
||| |||
Db 167 WDKSPTHP 174

Search completed: September 27, 2004, 12:27:47
Job time : 133 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 27, 2004, 12:09:05 ; Search time 39 Seconds
(without alignments)
27.131 Million cell updates/sec

Title: US-10-088-639A-2_COPY_88_98
Perfect score: 65
Sequence: 1 NSWSSGTHPV 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database : PIR 78:**

1: pir1:**
2: pir2:**
3: pir3:**
4: pir4:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	64.6	130	T43606	yopE chaperone syc
2	41	63.1	98	F97245	probable HD superf
3	41	63.1	339	T27931	hypothetical prote
4	41	63.1	456	B90543	hypothetical prote
5	40	61.5	18	C49254	Tcr C gamma 1 chai
6	40	61.5	110	C24775	T-cell receptor ga
7	40	61.5	120	B24775	T-cell receptor ga
8	40	61.5	132	S09713	Ig lambda chain V
9	40	61.5	629	T38214	hypothetical prote
10	40	61.5	914	T08081	probable myrosinas
11	40	61.5	988	T08102	myrosinase-binding
12	39	60.0	201	A88363	protein F26H11.1
13	39	60.0	270	T06528	lectin - garden pe
14	39	60.0	402	S74525	hypothetical prote
15	39	60.0	652	G85024	probable CHP-rich
16	39	60.0	995	T22942	hypothetical prote
17	39	60.0	1336	T02736	probable SCARECROW
18	38	58.5	295	E84862	hypothetical prote
19	38	58.5	451	S56818	hypothetical prote
20	38	58.5	476	D95318	Nife oxidoreductas
21	38	58.5	482	JC7332	endoglycosylceram
22	38	58.5	567	A97291	oligopeptide ABC t
23	38	58.5	1099	A59311	myosin VIII, ZMM3
24	38	58.5	1290	A56493	leucocyte common a
25	38	58.5	1369	D86178	hypothetical prote
26	38	58.5	1452	S17869	protein-tyrosine-p
27	38	58.5	1452	S17670	protein-tyrosine-p
28	38	58.5	1797	F69195	cell surface glyco
29	38	58.5	1898	S46216	leukocyte antigen-

30	37	56.9	108	2	S47184	Ig lambda chain -
31	37	56.9	109	2	S19663	Ig lambda chain V
32	37	56.9	125	2	C83704	hypothetical prote
33	37	56.9	271	2	B86918	hypothetical prote
34	37	56.9	294	2	JC5000	aryl sulfotransfer
35	37	56.9	295	2	JC5253	aryl sulfotransfer
36	37	56.9	295	2	S52791	phenol-sulfating p
37	37	56.9	295	2	I57945	aryl sulfotransfer
38	37	56.9	295	2	JC5248	aryl sulfotransfer
39	37	56.9	295	2	JC5249	aryl sulfotransfer
40	37	56.9	295	2	G01843	phenol sulfotransf
41	37	56.9	295	2	G02924	phenol sulfotransf
42	37	56.9	295	2	A55451	aryl sulfotransfer
43	37	56.9	295	2	S52794	aryl sulfotransfer
44	37	56.9	330	2	JC5935	exotoxose-related p
45	37	56.9	372	2	A53050	brefeldin A estera
46	37	56.9	375	1	WZWC6B	pectate lyase (EC
47	37	56.9	375	1	WZWC6C	pectate lyase (EC
48	37	56.9	375	2	S25262	pectate lyase (EC
49	37	56.9	394	2	D70550	probable fadB19 pr
50	37	56.9	666	1	VCMVHL	env polyprotein pr
51	37	56.9	1623	2	T01369	ABC transporter At
52	37	56.9	1732	2	E71442	hypothetical prote
53	37	56.9	2013	2	AD1129	probable peptidogl
54	37	56.9	2013	2	A11489	probable peptidogl
55	36	55.4	127	2	S70444	Ig lambda chain pr
56	36	55.4	150	2	H72658	hypothetical prote
57	36	55.4	190	2	S25740	Ig lambda chain -
58	36	55.4	234	2	A70736	hypothetical prote
59	36	55.4	250	2	S40405	MADS box protein o
60	36	55.4	336	2	JC4102	hypothetical 36.9k
61	36	55.4	394	2	AE1940	two-component resp
62	36	55.4	406	2	S57217	1-phosphatidylinos
63	36	55.4	426	2	B71408	probable acylamino
64	36	55.4	550	2	A12249	exopolysphatase
65	36	55.4	618	2	T49741	related to stress
66	36	55.4	620	2	E90575	lipoprotein limpor
67	36	55.4	665	1	VCWEM	env polyprotein -
68	36	55.4	689	2	AE1132	probable membrane
69	36	55.4	755	2	D75598	photoreceptor - De
70	36	55.4	811	2	S17908	hypothetical prote
71	36	55.4	831	2	S62751	probable DNA-direc
72	36	55.4	1036	2	B69368	hypothetical prote
73	36	55.4	1038	2	AG2187	hypothetical prote
74	35	53.8	8	2	A38887	T-cell receptor ga
75	35	53.8	96	2	S36060	Ig lambda chain -
76	35	53.8	110	2	S19672	Ig lambda chain V
77	35	53.8	115	2	S13726	Ig lambda chain V
78	35	53.8	161	2	C81711	conserved hypothet
79	35	53.8	223	2	AE2536	hypothetical prote
80	35	53.8	224	2	JC2075	2-haloacid dehalog
81	35	53.8	229	2	B38452	2-haloacid dehalog
82	35	53.8	237	2	A10871	thiol,disulfide in
83	35	53.8	259	2	S58988	cytochrome-c oxida
84	35	53.8	259	2	G95412	conserved hypothet
85	35	53.8	293	2	C89892	aspartate transcar
86	35	53.8	293	2	AF3393	ATPases of the pp
87	35	53.8	348	2	T02798	hypothetical prote
88	35	53.8	506	2	T01716	hypothetical prote
89	35	53.8	506	2	F85016	probable RING zinc
90	35	53.8	533	2	AG2328	2,3-bisphosphoglyc
91	35	53.8	536	2	T16458	hypothetical prote
92	35	53.8	552	2	A71864	ribosomal protein
93	35	53.8	556	2	G64569	ribosomal protein
94	35	53.8	556	2	S67097	probable membrane
95	35	53.8	603	2	B70886	probable penicilli
96	35	53.8	1053	2	T51016	related to CYC8 pr
97	35	53.8	1106	2	S38783	integrin alpha cha
98	35	53.8	1135	2	I61186	alpha-7 integrin -
99	34.5	53.1	335	2	S58892	signaling lymphocy
100	34.5	53.1	453	2	S32817	gasirin receptor -
101	34.5	53.1	559	1	JQ2010	transcription fact
102	34.5	53.1	602	2	S47880	NADH2 dehydrogenas

103 34.5 53.1 970 2 T28234 ORF MSV076 probabl
 104 34 52.3 107 2 JX0363 subtilisin-trypsin
 105 34 52.3 108 1 L3HUSH Ig lambda chain v-
 106 34 52.3 116 2 PC1268 subtilisin inhibit
 107 34 52.3 117 2 A34860 phospholipase A2 (
 108 34 52.3 118 2 B34860 phospholipase A2 (
 109 34 52.3 118 2 E34860 phospholipase A2 (
 110 34 52.3 118 2 C34860 phospholipase A2 (
 111 34 52.3 166 2 A57378 cyclin-dependent k
 112 34 52.3 166 2 B57378 cyclin-dependent k
 113 34 52.3 201 2 D71824 hypothetical prote
 114 34 52.3 206 2 A65189 yigW protein - Esc
 115 34 52.3 230 2 AF0005 tRNA (guanosine-2'
 116 34 52.3 232 2 AD1236 glycerophosphodias
 117 34 52.3 233 2 S25747 Ig lambda chain -
 118 34 52.3 233 2 A11598 glycerophosphodias
 119 34 52.3 235 2 B64453 probable formate a
 120 34 52.3 240 1 B64453 cytoplasmic Dnase
 121 34 52.3 260 2 A91225 probable deoxyribo
 122 34 52.3 260 2 AF0915 hypothetical prote
 123 34 52.3 260 2 G75301 hypothetical prote
 124 34 52.3 264 2 G86071 env polyprotein pr
 125 34 52.3 273 2 B31479 lectin-related sto
 126 34 52.3 290 2 S66355 aryl sulfotransfer
 127 34 52.3 295 2 S53399 glutamine PRPP am
 128 34 52.3 305 2 A69122 aspartate carbamoy
 129 34 52.3 307 2 D95148 aspartate carbamoy
 130 34 52.3 307 2 B98016 aspartate carbamoy
 131 34 52.3 310 2 B68625 lectectin - mouse
 132 34 52.3 313 2 JE0328 extensin precursor
 133 34 52.3 322 2 S23299 hypothetical prote
 134 34 52.3 325 2 E83165 cytochrome d ubiqu
 135 34 52.3 355 2 AF3604 knotted-like home
 136 34 52.3 397 2 T17008 knotted-like home
 137 34 52.3 398 2 T17009 probable prephenat
 138 34 52.3 413 2 T43017 prephenate dehydro
 139 34 52.3 431 2 T41005 GTP-binding protei
 140 34 52.3 459 2 T45911 Gi/S transition co
 141 34 52.3 465 2 D82804 probable porin PA4
 142 34 52.3 471 2 T02749 hypothetical prote
 143 34 52.3 484 2 A83082 hypothetical prote
 144 34 52.3 496 2 T51058 probable ABC trans
 145 34 52.3 516 2 S14694 F22M8.11 protein -
 146 34 52.3 530 2 C95268 hypothetical ABC t
 147 34 52.3 541 2 G86151 apolipoprotein B -
 148 34 52.3 578 2 AF0599 hypothetical prote
 149 34 52.3 596 2 S32802
 150 34 52.3 681 2 T19429

ALIGNMENTS

RESULT 1
 T43606
 yopE Chapterone syce - Yersinia pestis plasmid pCD1
 C:Species: Yersinia pestis
 C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000
 R:Hu, P.; Elliott, J.; McCready, P.; Skowronski, E.; Barnes, J.; Kobayashi, A.; Brubaker
 J. Bacteriol. 180, 5192-5202, 1998
 A:Title: Structural organization of virulence-associated plasmids of Yersinia pestis.
 A:Reference number: 222578; MUID:98422474; PMID:9748454
 A:Accession: T43606
 A:Status: preliminary; translated from GB/EMBL/DBSJ
 A:Molecule type: DNA
 A:Residues: 1-130 <HUP>
 A:Cross-references: EMBL:AF053946; NID:g2996222; PIDN:AC62588.1; PID:g2996265
 A:Experimental source: strain KIM
 R:Perry, R.D.; Straley, S.C.; Fetherston, J.D.; Rose, D.J.; Gregor, J.; Blattner, F.R.
 Infect. Immun. 66, 4611-4623, 1998
 A:Title: DNA sequencing and analysis of the low-Ca2+-response plasmid pCD1 of Yersinia P
 A:Reference number: 222273; MUID:98427122; PMID:9746557

A:Accession: T42911
 A:Status: preliminary; translated from GB/EMBL/DBSJ
 A:Molecule type: DNA
 A:Residues: 1-130 <PER>
 A:Cross-references: EMBL:AF074612; NID:g3822037; PIDN:AAC69819.1; PID:g3822099
 A:Experimental source: strain KIM5
 C:Genetics:
 A:Gene: syce
 A:Genome: plasmid pCD1
 C:Superfamily: Yersinia pestis plasmid pCD1 hypothetical protein syce
 Query Match 64.6%; Score 42; DB 2; Length 130;
 Best Local Similarity 70.0%; Pred. No. 4.2;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 2 SWDSSGTHPV 11
 ||| ||||
 Db 79 SWDEVGGHPV 88
 ||| ||||
 RESULT 2
 F97245
 probable HD superfamily hydrolase [imported] - Clostridium acetobutylicum
 C:Species: Clostridium acetobutylicum
 C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
 C:Accession: F97245
 R:Nolling, J.; Breton, G.; Onelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
 J. Bacteriol. 183, 4823-4838, 2001
 A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
 A:Reference number: A96900; MUID:21359325; PMID:21359325
 A:Accession: F97245
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-98 <KUR>
 A:Cross-references: GB:AE001437; PIDN:AAK80753.1; PID:GI5025850; GSPDB:GNO0168
 A:Experimental source: Clostridium acetobutylicum ATCC824
 C:Genetics:
 A:Gene: CAC2809
 Query Match 63.1%; Score 41; DB 2; Length 98;
 Best Local Similarity 77.8%; Pred. No. 4.6;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 2 SWDSSGTHPV 10
 ||| ||||
 Db 40 SWDQSGTTP 48
 ||| ||||

RESULT 3
 T27931
 hypothetical protein ZK596.3 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T27931
 R:McMurray, A.
 submitted to the EMBL Data Library, February 1996
 A:Reference number: Z20441
 A:Accession: T27931
 A:Status: preliminary; translated from GB/EMBL/DBSJ
 A:Molecule type: DNA
 A:Residues: 1-339 <WIL>
 A:Cross-references: EMBL:Z69386; PIDN:CAA93432.1; GSPDB:GNO00029; CESP:ZK596.3
 A:Experimental source: clone ZK596
 C:Genetics:
 A:Gene: CESP:ZK596.3
 A:Map position: 4
 A:Introns: 62/3; 110/3; 134/3; 200/1; 283/1
 Query Match 63.1%; Score 41; DB 2; Length 339;
 Best Local Similarity 70.0%; Pred. No. 18;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 SWDSSGTHPV 11
||:|||||
Db 48 SWNSNGTHV 57

RESULT 4

B90543
Hypothetical protein MYPV 2500 [imported] - Mycoplasma pulmonis (strain UAB CTIP)
C:Species: Mycoplasma pulmonis
C>Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 03-Aug-2001
C:Accession: B90543
R:Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer, I.;
Nucleic Acids Res. 29, 2145-2153, 2001
A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulm
A:Reference number: A99512; MUID:21267165; PMID:11353084
A:Accession: B90543
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-456 <KUR>
A:CROSS-references: GB:AL445566; PID:G14089663; PIDN:CAC13423.1; GSPDB:GN00153
A:Experimental source: strain UAB CTIP
C:Genetics:
A:Gene: MYPV 2500
A:Genetic code: SGC3

Query Match 63.1%; Score 41; DB 2; Length 456;
Best Local Similarity 70.0%; Pred. No. 24;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 NSWDSSGTHP 10
|||||
Db 19 NSWDISKTP 28

RESULT 5

C49254
TCR C gamma 1 chain V-J region - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999
C:Accession: C49254
R:Ezquerria, A.; Wilde, D.B.; McConnell, T.J.; Sturmhofel, K.; Valas, R.B.; Shevach, E.M.;
Eur. J. Immunol. 22, 491-498, 1992
A:Title: Mouse autoreactive gamma/delta T cells. II. Molecular characterization of the T
A:Reference number: A49037; MUID:92164730; PMID:1311262
A:Accession: C49254
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-18 <EQZ>
A:CROSS-references: GB:S90689; NID:G246310; PIDN:AAB21558.1; PID:G246311
A:Experimental source: FF3 cells
A:Note: sequence extracted from NCBI backbone (NCBIN:90689, NCBIP:90691)

Query Match 61.5%; Score 40; DB 2; Length 18;
Best Local Similarity 77.8%; Pred. No. 1.1;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 WDSSGTHPV 11
|||||
Db 8 WDSSGFHKV 16

RESULT 6

C24775
T-cell receptor gamma chain precursor V region (V4) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 05-Jun-1988 #sequence_revision 05-Jun-1988 #text_change 24-Sep-1999
C:Accession: C24775
R:Garman, R.D.; Doherty, P.J.; Raullet, D.H.
Cell 45, 733-742, 1986
A:Title: Diversity, rearrangement, and expression of murine T cell gamma genes.
A:Reference number: A90880; MUID:86218086; PMID:3486721
A:Accession: C24775
A:Molecule type: DNA

A:Residues: 1-110 <GAR>
A:Note: this sequence was determined from the differentiated gene
A:Note: the authors translated the codon ACG for residue 2 as Ser and ACA for residue 59
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: T-cell receptor

Query Match 61.5%; Score 40; DB 2; Length 110;
Best Local Similarity 77.8%; Pred. No. 7.7;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 WDSSGTHPV 11
|||||
Db 102 WDSSGFHKV 110

RESULT 7

B24775
T-cell receptor gamma chain precursor V region (V3) - mouse
C:Species: Mus musculus (house mouse)
C>Date: 05-Jun-1988 #sequence_revision 05-Jun-1988 #text_change 23-Jul-1999
C:Accession: B24775
R:Garman, R.D.; Doherty, P.J.; Raullet, D.H.
Cell 45, 733-742, 1986
A:Title: Diversity, rearrangement, and expression of murine T cell gamma genes.
A:Reference number: A90880; MUID:86218086; PMID:3486721
A:Accession: B24775
A:Molecule type: DNA
A:Residues: 1-120 <GAR>
A:CROSS-references: GB:M1337; NID:G201668; PIDN:AAA40316.1; PID:G201669
A:Note: this sequence was determined from the differentiated gene
A:Note: the authors' translation contained an addition Cys after 14-Val
C:Genetics:
A:Introns: 15/1
A:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: T-cell receptor

Query Match 61.5%; Score 40; DB 2; Length 120;
Best Local Similarity 77.8%; Pred. No. 8.5;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 WDSSGTHPV 11
|||||
Db 112 WDSSGFHKV 120

RESULT 8

S09713
Ig lambda chain V region - human
C:Species: Homo sapiens (man)
C>Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 21-Jan-2000
C:Accession: S09713
R:Hughes-Jones, N.C.; Bye, J.M.; Beale, D.; Coadwell, J.
Biochem. J. 268, 135-140, 1990
A:Title: Nucleotide sequences and three-dimensional modelling of the VH and VL domains of
A:Reference number: S09710; MUID:90262535; PMID:2111699
A:Accession: S09713
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-132 <HUG>
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F;34-110/Domain: immunoglobulin homology <IMM>

Query Match 61.5%; Score 40; DB 2; Length 132;
Best Local Similarity 75.0%; Pred. No. 9.4;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 WDSSGTHP 10
|||||
Db 111 WDSSSAHP 118

RESULT 9

T38214
 hypothetical protein SPAC22H10.11c - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 04-Mar-2000
 C:Accession: T38214
 R:Grishchuk, K.; McIntosh, J.R.; Devlin, K.; Churcher, C.; Barrell, B.G.; Rajandream, M.
 submitted to the EMBL Data Library, February 1996
 A:Reference number: Z21778
 A:Accession: T38214
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-629 <DEV>
 A:Cross-references: EMBL:Z69730; PIDN:CAA93611.1; GSPDB:GN00066; SPDB:SPAC22H10.11c
 A:Experimental source: strain 972h; cosmid c22H10
 C:Genetics:
 A:Gene: SPDB:SPAC22H10.11c
 A:Map position: 1
 C:Superfamily: Schizosaccharomyces pombe hypothetical protein SPAC22H10.11c

Query Match 61.5%; Score 40; DB 2; Length 629;
 Best Local Similarity 60.0%; Pred. No. 51;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 NSWDSSGTHP 10
 |||||:|
 Db 100 NSWKXSSHP 109

RESULT 10
 T08081
 probable myrosinase-binding protein - rape
 N:Alternate names: jasmonate inducible protein
 C:Species: Brassica napus (rape)
 C:Date: 21-May-1999 #sequence_revision 21-May-1999 #text_change 11-May-2000
 C:Accession: T08081
 R:Geshi, N.; Brandt, A.
 Planta 204, 295-304, 1998
 A:Title: Two jasmonate-inducible myrosinase-binding proteins from Brassica napus L. seed
 A:Reference number: Z16340; MUID:98192006; PMID:9530873
 A:Accession: T08081
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-914 <GES>
 A:Cross-references: EMBL:Y11482; NID:e1023100; PIDN:CAA72270.1; PID:e304149
 A:Experimental source: cv. Global; isolate a4; seedlings
 A:Note: Jasmonate inducible

Query Match 61.5%; Score 40; DB 2; Length 914;
 Best Local Similarity 66.7%; Pred. No. 77;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 NSWDSSGTHP 9
 |||||:|
 Db 683 NQWDDSGDH 691

RESULT 11
 T08102
 myrosinase-binding protein (clone indmbp2) - rape
 C:Species: Brassica napus (rape)
 C:Date: 21-May-1999 #sequence_revision 21-May-1999 #text_change 21-Jul-2000
 C:Accession: T08102
 R:Taipalensuu, J.; Eriksson, S.J.; Rask, L.
 Eur. J. Biochem. 250, 680-688, 1997
 A:Title: The myrosinase binding protein from Brassica napus seeds possesses lectin activ
 A:Reference number: Z16355; MUID:98121188; PMID:9461290
 A:Accession: T08102
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-988 <TAI>
 A:Cross-references: EMBL:Y09437; NID:g1711295; PIDN:CAA70587.1; PID:g1711296
 A:Experimental source: cv. Hanna; leaves

Query Match 61.5%; Score 40; DB 2; Length 988;
 Best Local Similarity 66.7%; Pred. No. 83;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 NSWDSSGTHP 9
 |||||:|
 Db 699 NQWDDSGDH 707

RESULT 12
 A88363
 protein F26H11.1 [imported] - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
 C:Accession: A88363
 R:anonymous, The C. elegans Sequencing Consortium.
 Science 282, 2012-2018, 1998
 A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
 A:Reference number: A75000; MUID:99069613; PMID:9851916
 A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_eleg
 A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
 A:Accession: A88363
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-201 <STO>
 A:Cross-references: GB:chr_II; PIDN:CA804194.1; PID:g3876448; GSPDB:GN00020; CESP:F26H11.1
 A:Note: Predicted using GeneFinder
 C:Genetics:
 A:Gene: F26H11.1
 A:Map position: 2

Query Match 60.0%; Score 39; DB 2; Length 201;
 Best Local Similarity 66.7%; Pred. No. 22;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 SWDSSGTHP 10
 |||||:|
 Db 101 SWSNSGFHP 109

RESULT 13
 T06528
 lectin - garden pea
 C:Species: Pisum sativum (garden pea)
 C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 20-Aug-1999
 C:Accession: T06528
 R:Kardailsky, I.V.; Sherrier, D.J.; Brewin, N.J.
 Plant Physiol. 111, 49-60, 1995
 A:Title: Identification of a new pea gene PnLec1 that encodes a lectin-like glycoprotein
 A:Reference number: Z15736
 A:Accession: T06528
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-270 <KAR>
 A:Cross-references: EMBL:U31981; NID:g974303; PIDN:AAC49367.1; PID:g974304
 A:Experimental source: cv. Wisconsin Perfection
 C:Genetics:
 A:Gene: Nlec1
 C:Superfamily: plant lectin

Query Match 60.0%; Score 39; DB 2; Length 270;
 Best Local Similarity 66.7%; Pred. No. 30;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 NSWDSSGTHP 9
 |||||:|
 Db 162 NSWDPEGRH 170

RESULT 14
 S74525
 hypothetical protein slr0731 - Synechocystis sp. (strain PCC 6803)
 C:Species: Synechocystis sp.

A:Variety: PCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C:Accession: S74525
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O.K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* s.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S74525
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-402 <RAN>
A:Cross-references: EMBL:D90899; GB:AB001339; NID:g1651650; PIDN:BAAL6677.1; PID:g165174
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Superfamily: *Synechocystis* hypothetical protein slr0731

Query Match 60.0%; Score 39; DB 2; Length 402;
Best Local Similarity 66.7%; Pred. No. 47;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NSWDSSGTH 9
Db 197 DSWESSSTH 205
:|||||

RESULT 15
G85024
probable CHP-rich zinc finger protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001
C:Accession: G85024
R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring Nature 402, 769-777, 1999
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: A85001; MUID:20083488; PMID:10617198
A:Accession: G85024
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-652 <STO>
A:Cross-references: GB:NC_001268; NID:g7268577; PIDN:CAB80686.1; GSPDB:GN00140
C:Genetics:
A:Gene: AT4G01930
A:Map position: 4

Query Match 60.0%; Score 39; DB 2; Length 652;
Best Local Similarity 54.5%; Pred. No. 79;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 NSWDSSGTHPV 11
Db 393 NKWCSACTHPI 403
:|||||

RESULT 16
T22942
hypothetical protein F58G11.2 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T22942
R:Percy, C.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z19640
A:Accession: T22942
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-995 <WIL>
A:Cross-references: EMBL:Z81094; PIDN:CAB03153.1; GSPDB:GN00023; CESP:F58G11.2
A:Experimental source: clone F58G11
C:Genetics:
A:Gene: CESP:F58G11.2
A:Map position: 5
A:Introns: 79/1; 195/1; 642/1; 678/2; 796/1

Query Match 60.0%; Score 39; DB 2; Length 995;
Best Local Similarity 66.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 SWDSSGTHP 10
Db 382 SWTNSGLHP 390
:|||||

RESULT 17
T02736
probable SCARECROW gene regulator [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 02-Feb-2001
C:Accession: T02736; A84692
R:Rounsley, S.D.; Lin, X.; Kaul, S.; Shea, T.P.; Fujii, C.Y.; Mason, T.M.; Shen, M.; Rong submitted to the EMBL Data Library, August 1998
A:Description: Arabidopsis thaliana chromosome II BAC T914 genomic sequence.
A:Reference number: Z14710
A:Accession: T02736
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1336 <ROU>
A:Cross-references: EMBL:AC005315; NID:g3461834; PID:g3461846
A:Experimental source: cultivar Columbia
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; Matus, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: A84692
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1336 <STO>
A:Cross-references: GB:AE002093; NID:g3461846; PIDN:AAC33232.1; GSPDB:GN00139
C:Genetics:
A:Gene: AT2G29060; T914.14
A:Map position: 2
A:Introns: 694/2

Query Match 60.0%; Score 39; DB 2; Length 1336;
Best Local Similarity 75.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSWDSSGT 8
Db 798 NSWDASGS 805
:|||||

RESULT 18
E84862
hypothetical protein At2g43140 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: E84862
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; Matus, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: E84862
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-295 <STO>
A:Cross-references: GB:AE002093; NID:g3763923; PIDN:AAC64303.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g43140
A:Map position: 2

Best Local Similarity 55.6%; Pred. No. 50;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 NSWDSSGTH 9
|:|:|:|:
Db 172 NNWDNSSH 180

RESULT 19

S56818
hypothetical protein YJL046w - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein J1171
C:Species: Saccharomyces cerevisiae
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 19-Apr-2002
C:Accession: S56818
R:Pohl, T.M.; Aljinovic, G.
A:Title: Cloning and expression of gene encoding a novel endoglycosylceramidase of Rhodococcus
submitted to the Protein Sequence Database, September 1995
A:Reference number: S56793
A:Accession: S56818
A:Molecule type: DNA
A:Residues: 1-451 <TO>
A:Cross-references: EMBL:Z49321; NID:g1008176; PIDN:CAA89337.1; PID:g1008177; GSPDB:GN00
C:Genetics:
A:Gene: MIPS:YJL046w
A:Cross-references: SGD:S0003582
A:Map position: 10L
C:Superfamily: lipoate-protein ligase

Query Match 58.5%; Score 38; DB 1; Length 451;
Best Local Similarity 66.7%; Pred. No. 79;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 WDSSGTHPV 11
|:|:|:|:
Db 333 WSSGVHVS 341

RESULT 20

D95318
NifE oxidoreductase nife [imported] - Sinorhizobium meliloti (strain 1021) magaplasmid p
C:Species: Sinorhizobium meliloti
C>Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C:Accession: D95318
R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows
; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti
A:Reference number: A95262; MUID:21396509; PMID:11481432
A:Accession: D95318
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-476 <KUR>
A:Cross-references: GB:AE006469; PIDN:AAK65110.1; PID:g14523548; GSPDB:GN00165
A:Experimental source: strain 1021, megaplasmid pSymA
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hymen, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: nife

C:Superfamily: dinitrogenase alpha chain; nitrogenase vanadium-iron protein alpha chain

Query Match 58.5%; Score 38; DB 2; Length 476;
Best Local Similarity 75.0%; Pred. No. 84;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NSWDSSGT 8

Db 73 NSWDNRGT 80

RESULT 21

JC7332
endoglycosylceramidase (EC 3.2.1.123) - Rhodococcus sp.
N:Alternate names: Cer glycanase; endoglycosylceramidase
C:Species: Rhodococcus sp.
C>Date: 08-Sep-2000 #sequence_revision 08-Sep-2000 #text_change 08-Sep-2000
C:Accession: JC7332
R:Sakaguchi, K.; Okino, N.; Sueyoshi, N.; Izu, H.; Ito, M.
J. Biochem. 128, 145-152, 2000
A:Title: Cloning and expression of gene encoding a novel endoglycosylceramidase of Rhodococcus
submitted to GenBank, May 1999
A:Reference number: JC7332
A:Accession: JC7332
A:Molecule type: DNA
A:Residues: 1-482 <SAK>
A:Experimental source: strain C9
C:Comment: This enzyme, a member of type II enzyme, cleaves the glycosidic linkage between
C:Genetics:
A:Start codon: GTG
A:Gene: EGCCase
C:Keywords: glycosidase; hydrolase

Query Match 58.5%; Score 38; DB 2; Length 482;
Best Local Similarity 50.0%; Pred. No. 85;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 NSWDSSGTHP 10
|:|:|:|:
Db 184 NFWNTTGKHP 193

RESULT 22

A97291
oligopeptide ABC transporter, periplasmic binding component CAC3179 [imported] - Clostridium
C:Species: Clostridium acetobutylicum
C>Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
C:Accession: A97291
R:Nolling, J.; Bretton, G.; Onelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 193, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: A97291
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-567 <KUR>
A:Cross-references: GB:AE001437; PIDN:AAK81116.1; PID:g15026248; GSPDB:GN00168
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
A:Gene: CAC3179

Query Match 58.5%; Score 38; DB 2; Length 567;
Best Local Similarity 45.5%; Pred. No. 1e+02;
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 NSWDSSGTHPV 11
|:|:|:|:
Db 81 NGWDGNATDPI 91

RESULT 23

A59311
myosin VIII, ZMM3 - maize (fragment)
C:Species: Zea mays (maize)
C>Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 08-Sep-2000
C:Accession: A59311
R:Li, L.; Pesacreta, T.C.
submitted to GenBank, May 1999
A:Reference number: A59311
A:Accession: A59311

A;Status: preliminary; not compared with conceptual translation
 A:Molecule type: mRNA
 A;Residues: 1-1099 <LIU>
 A;Cross-references: GB:AF147738; NID:g4885025; PIDN:AAD31926.1; PID:g4885026
 A;Experimental source: cultivar B73 inbred line
 C;Genetics:

A;Gene: zmm3
 C;Superfamily: myosin heavy chain 95F; myosin motor domain homology
 F;101-757/Domain: myosin motor domain homology <MMO>

Query Match 58.5%; Score 38; DB 2; Length 1099;
 Best Local Similarity 60.0%; Pred. No. 2.1e+02;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 SWDSSGTHPV 11

|||||

Db 976 SWESNGNHV 985

RESULT 24

A56493

leucocyte common antigen-related protein (LAR) - rat (fragment)

N;Alternate names: LAR receptor-linked tyrosine phosphatase

N;Contains: protein-tyrosine-phosphatase (EC 3.1.3.48)

C;Species: Rattus norvegicus (Norway rat)

C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 21-Jan-2000

C;Accession: A56493; I55393

R;Zhang, J.S.; Longo, F.M.

J. Cell Biol. 128, 415-431, 1995

A;Title: LAR tyrosine phosphatase receptor: alternative splicing is preferential to the

A;Reference number: A56493; MUID:95146548; PMID:7844155

A;Accession: A56493

A;Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A;Residues: 1-1290 <RES>

A;Cross-references: EMBL:X83505; NID:g732918; PIDN:CAA58495.1; PID:g732919

R;Longo, F.M.; Martignetti, J.A.; Le Beau, J.M.; Zhang, J.S.; Barnes, J.P.; Brosius, J.

J. Biol. Chem. 268, 26503-26511, 1993

A;Title: Leukocyte common antigen-related receptor-linked tyrosine phosphatase. Regulation

A;Reference number: I55393; MUID:94075340; PMID:8253779

A;Accession: I55393

A;Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A;Residues: 831-1290 <RE2>

A;Cross-references: EMBL:U00477; NID:g392565; PIDN:AAC04306.1; PID:g392566

C;Superfamily: leukocyte antigen-related protein; fibronectin type III repeat homology;

ogy

C;Keywords: phosphoprotein; phosphoric monoester hydrolase; tyrosine-specific phosphatase

F;80-166/Domain: fibronectin type III repeat homology <3PR>

F;667-1290/Domain: leukocyte common antigen cytosolic domain homology <LAC>

F;1047-1270/Domain: protein-tyrosine-phosphatase homology <PTP2>

F;931/Active site: Cys (phosphocysteine intermediate) #status predicted

F;937/Binding site: substrate phosphate (Arg) #status predicted

F;1222/Active site: Cys (phosphocysteine intermediate) #status predicted

F;1228/Binding site: substrate phosphate (Arg) #status predicted

Query Match 58.5%; Score 38; DB 2; Length 1290;

Best Local Similarity 60.0%; Pred. No. 2.5e+02;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 SWDSSGTHPV 11

|||||

Db 5 TWDSGNTEPV 14

RESULT 25

D86178

hypothetical protein [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 20-Apr-2001

C;Accession: D86178

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: D86178

A;Status: preliminary

A:Molecule type: DNA

A;Residues: 1-1369 <STO>

A;Cross-references: GB:AE005172; NID:g2494131; PIDN:AAB80640.1; GSPDB:GN00141

C;Genetics:

A;Map position: 1

C;Superfamily: xanthine dehydrogenase; ferredoxin [2Fe-2S] homology

C;Keywords: 2Fe-2S; metalloprotein

F;43,48,51,73/Binding site: 2Fe-2S cluster (Cys) (covalent) #status predicted

Query Match 58.5%; Score 38; DB 2; Length 1369;

Best Local Similarity 75.0%; Pred. No. 2.6e+02;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 SWDSSGTH 9

|||||

Db 531 SWDSEGH 538

Search completed: September 27, 2004, 12:16:38

Job time : 44 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2005, 17:57:45 ; Search time 82.8438 seconds
(without alignments)
51.354 Million cell updates/sec

Title: US-10-088-639A-2_COPY_88_98
Perfect score: 65
Sequence: 1 NSWDSSTHPV 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	65	100.0	249	4	AAB68087
2	51	78.5	11	5	AAO15005
3	49	75.4	110	6	ABR54947
4	46	70.8	249	5	ABP45104
5	46	70.8	249	7	ADG95931
6	44	67.7	247	5	ABP45105
7	44	67.7	247	7	ADG95932
8	43	66.2	245	8	ADSO9248
9	43	66.2	508	5	ABB06006
10	41	63.1	246	5	ABP45969
11	41	63.1	246	7	ADG96796
12	40	61.5	13	5	AAU70345
13	40	61.5	107	6	ADBO9502
14	40	61.5	110	2	AAW15538
15	40	61.5	111	2	AAW12264
16	40	61.5	112	6	ABR82488
17	40	61.5	120	6	ABR82487
18	40	61.5	249	5	ABP44529
19	40	61.5	249	5	ABP44342
20	40	61.5	249	7	ADG95169
21	40	61.5	249	7	ADG95356
22	40	61.5	448	7	ABO75320
23	40	61.5	469	5	ABP65370
24	39	60.0	61	4	ABG01449
25	39	60.0	86	8	ABO54515

26	39	60.0	162	4	AAU28269
27	39	60.0	305	5	AAU75568
28	39	60.0	305	6	ABJ39144
29	39	60.0	305	7	ADG59481
30	39	60.0	305	7	ADM04464
31	39	60.0	311	4	AAW1910
32	39	60.0	324	4	AAU28081
33	39	60.0	373	8	ADG32664
34	39	60.0	683	6	ABU03132
35	39	60.0	995	8	ADN23915
36	39	60.0	1336	5	AAU92973
37	39	60.0	1336	6	ADA15573
38	39	60.0	1336	7	ADD30544
39	39	60.0	1336	7	ADG31525
40	39	60.0	1336	8	ADI41741
41	39	60.0	1336	8	ADO02241
42	38	58.5	11	8	ADR38721
43	38	58.5	53	5	ABP01278
44	38	58.5	92	2	AAW72861
45	38	58.5	111	7	ADM05567

ALIGNMENTS

RESULT 1

AAB68087
ID AAB68087 standard; protein; 249 AA.

AC AAB68087;

XX 09-JUL-2001 (first entry)

DE An anti-alpha6beta4 integrin light chain linked to a heavy chain.

XX Gastrointestinal epithelial tumour cell; alpha6beta4 integrin;

KW tumour-associated antigen; metastatic disease; malignant disease;

KW tumour typing; tumour screening; tumour.

XX Synthetic.

OS Macaca fascicularis.

XX Key Location/Qualifiers
Region 23..33

FT /note= "Complementarity determining region (CDR) 1 of the light chain"

FT /note= "Complementarity determining region (CDR) 2 of the light chain"

FT /note= "Complementarity determining region (CDR) 3 of the light chain"

FT /note= "linker"

FT /note= "Complementarity determining region (CDR) 1 of the heavy chain"

FT /note= "Complementarity determining region (CDR) 1 of the heavy chain"

FT /note= "Complementarity determining region (CDR) 1 of the heavy chain"

FT /note= "Complementarity determining region (CDR) 1 of the heavy chain"

FT /note= "Complementarity determining region (CDR) 1 of the heavy chain"

XX WO200130854-A2.

PN 03-MAY-2001.

PD 26-OCT-2000; 2000WO-SB002082.

XX 28-OCT-1999; 99SE-00003895.

XX (ACTI-) ACTIVE BIOTECH AB.

XX Brodin TN, Karlstroem PJ, Ohlsson LG, Tordsson MJ, Kearney PP;
 PI Nilson BHK;
 XX WPI; 2001-308619/32.
 DR N-PSDB; AAF84797.
 XX Novel antibody for diagnosis, treatment of human metastatic and malignant
 PT diseases, has binding structure for target structure displayed on cell
 PT surface of human gastrointestinal epithelial tumor and normal cells.
 XX Claim 1; Page 55-56; 75pp; English.
 XX The present sequence represents a Monkey antibody light chain linked to a
 CC heavy chain. The antibody binds to a target structure displayed in and on
 CC the cell surface of human gastrointestinal epithelial tumour cells and in
 CC a subpopulation of normal human gastrointestinal epithelial cells. The
 CC target structure especially comprises alpha6beta4 integrin. This is a
 CC tumour-associated antigen. The antibody, and its fragments, are useful
 CC for treating conditions based on anti-angiogenic mechanism in humans.
 CC They are useful for treating human metastatic and malignant disease, for
 CC in vitro, in vivo diagnosis and prognosis of human malignant disease,
 CC comprising tumour typing, tumour screening, tumour diagnosis and
 CC prognosis and monitoring premalignant conditions. Quantitative in vivo
 CC diagnosis is carried out by determining the localization of antibody to
 CC tumour deposits in humans
 XX Sequence 249 AA;
 SQ
 Query Match 100.0%; Score 65; DB 4; Length 249;
 Best Local Similarity 100.0%; Pred. No. 0.014;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 NSWDSSGTHPV 11
 Db 88 NSWDSSGTHPV 98
 RESULT 2
 AA015005
 ID AA015005 standard; peptide; 11 AA.
 XX
 AC AA015005;
 XX
 DT 16-AUG-2002 (first entry)
 XX
 DE GPI-linked cell surface receptor-specific VL CDR3 peptide sequence.
 XX
 KW Ribosome display system; specific binding pair; sbp; anti-FITC; VL CDR3;
 KW scFv antibody generation; GPI-linked cell surface receptor.
 XX
 OS Unidentified.
 XX
 PN WO200175097-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 26-MAR-2001; 2001WO-GB001319.
 XX
 PR 31-MAR-2000; 2000US-0193802P.
 XX
 PA (CAMB-) CAMBRIDGE ANTIBODY TECHNOLOGY.
 XX
 PI Osbourn J, Holet T;
 XX
 DR WPI; 2002-010788/01.
 XX
 PT Obtaining specific binding pair member (I) that binds a complementary
 PT specific binding pair member of encapsidating specific binding pair
 PT member/ribosome complexes in a viral coat, comprises the use of a
 PT ribosome display system.
 XX
 PS Example 7; Page 42; 61pp; English.

XX The invention comprises a method (ribosome display system) for obtaining
 CC a specific binding pair (sbp) member that binds a complementary sbp
 CC member of interest. The method involves incubating mRNA molecules
 CC encoding an sbp and lacking an in-frame stop codon, allowing ribosome
 CC translation of the mRNA to produce the encoded sbp member, forming
 CC complexes comprising ribosome, mRNA, and encoded sbp member displayed on
 CC the ribosome. The ribosome display system is useful for the selection of
 CC an sbp member able to bind a complementary sbp member. The present amino
 CC acid sequence represents a GPI-linked cell surface receptor-specific VL
 CC CDR3 peptide
 XX Sequence 11 AA;
 SQ
 Query Match 78.5%; Score 51; DB 5; Length 11;
 Best Local Similarity 81.8%; Pred. No. 0.1;
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 NSWDSSGTHPV 11
 Db 1 NSWDSSGNHVV 11
 RESULT 3
 ABR54947
 ID ABR54947 standard; protein; 110 AA.
 XX
 AC ABR54947;
 XX
 DT 30-JUN-2003 (first entry)
 XX
 DE IgG light chain clone HBL4a 3D6 SEQ ID NO:173.
 XX
 KW Engineered template; single primer amplification; antibody library;
 KW nucleic acid amplification.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 PN WO2003025202-A2.
 XX
 PD 27-MAR-2003.
 XX
 PF 19-SEP-2002; 2002WO-US029889.
 XX
 PR 19-SEP-2001; 2001US-0323455P.
 XX
 PA (ALEX-) ALEXION PHARM INC.
 XX
 PI Bowdish KS, Frederickson S, Maruyama T, Lin Y, Renshaw M;
 XX
 DR WPI; 2003-313359/30.
 XX
 PT Amplifying nucleic acid by contacting engineered nucleic acid strand
 PT having predetermined sequence at one end and sequence complementary to
 PT predetermined sequence at other end, with primer having predetermined
 PT sequence.
 XX
 PS Example 3; Fig 8e; 68pp; English.
 XX
 CC The present invention describes a method (M1) for amplifying a nucleic
 CC acid strand. M1 comprises providing an engineered nucleic acid strand (S)
 CC having a predetermined sequence at one end and a sequence complementary
 CC to the predetermined sequence at the other, and contacting (S) with a
 CC primer having the predetermined sequence in the presence of a polymers
 CC and nucleotides under conditions suitable for polymerisation of the
 CC nucleotides. Also described is an engineered nucleic acid strand (I)
 CC having a predetermined sequence at one end and a sequence complementary
 CC to the predetermined sequence at the other end. M1 is useful for
 CC amplifying a nucleic acid. M1 can be used for producing an antibody
 CC library. M1 is useful for preparing amplified products that can be
 CC ligated into a suitable expression vector, where the vector can be used
 CC to transform an appropriate host organism to produce the polypeptide or

CC protein encoded by the target sequence. M1 is useful to amplify a family
 CC of related sequences to build a complex library such as, for example an
 CC antibody library. M1 is useful not only for producing large amounts of
 CC one target nucleic acid sequence, but also for amplifying simultaneously
 CC more than one different target nucleic acid sequence located on the same
 CC or different nucleic acid molecules. ACC62635 to ACC62753 and ABR54841 to
 CC ABR54998 represent sequence used in the exemplification of the present
 CC invention

XX SQ Sequence 110 AA;

Query Match 75.4%; Score 49; DB 6; Length 110;
 Best Local Similarity 88.9%; Pred. No. 2.7;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 3 WDSGTHPV 11
 |||||
 Db 92 WDSGDHPV 100

RESULT 4

ID ABP45104 standard; protein; 249 AA.

XX AC ABP45104;

XX DT 19-AUG-2002 (first entry)

XX DE Human BlyS binding scFv SEQ ID 1115.

XX BlyS; B lymphocyte stimulator; TNF superfamily; human; cytostatic;
 KW tumour necrosis factor; B cell proliferation; B cell differentiation;
 KW immunosuppressive; immunostimulant; immunomodulatory; antirheumatic;
 KW antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;
 KW systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;
 KW common variable immunodeficiency; acquired immunodeficiency syndrome.

XX OS Homo sapiens.

XX PN WO200202641-A1.

XX PD 10-JAN-2002.

XX PF 15-JUN-2001; 2001WO-US019110.

XX PR 16-JUN-2000; 2000US-0212210P.

XX PR 17-OCT-2000; 2000US-0240816P.

XX PR 16-MAR-2001; 2001US-0276248P.

XX PR 21-MAR-2001; 2001US-027379P.

XX PR 25-MAY-2001; 2001US-0293499P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PA (CAMB-) CAMBRIDGE ANTIBODY TECHNOLOGY.

XX PI Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;

XX DR WPI; 2002-114799/15.

XX Antibodies against B Lymphocyte Stimulating polypeptides, useful for the
 PT diagnosis and treatment of cancers and immune disorders.

XX Claim 1; Page 1733-1734; 3148pp; English.

CC This invention describes novel antibodies that immunospecifically bind to
 CC B lymphocyte stimulator (BlyS) polypeptides. BlyS is a member of the
 CC tumour necrosis factor (TNF) super family and induces B cell
 CC proliferation and differentiation. The antibodies of the invention have
 CC cytostatic, immunosuppressive, immunostimulant, immunomodulatory,
 CC antirheumatic and antiAIDS activity and can be used in vaccines to
 CC inhibit the expression and activity of BlyS. The antibodies bind to BlyS
 CC and so may be used to detect and quantitate the presence of BlyS in
 CC biological samples and may be used in this way to diagnose disease
 CC associated with aberrant expression of BlyS. They may also be

CC administered to treat diseases associated with aberrant BlyS expression
 CC and activity such as cancer, immune, and autoimmune disorders and
 CC diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis,
 CC immunodeficiency (e.g. common variable immunodeficiency (CVID) and
 CC acquired immunodeficiency syndrome (AIDS)). ABP43990-ABP47228 represent
 CC the antibodies and fragments of the antibodies described in the method of
 CC the invention

XX SQ Sequence 249 AA;

Query Match 70.8%; Score 46; DB 5; Length 249;
 Best Local Similarity 72.7%; Pred. No. 21;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 1 NSWSSSGTHPV 11

Db 228 HSWSSGNHV 238

RESULT 5

ID ADG95931 standard; protein; 249 AA.

XX AC ADG95931;

XX DT 11-MAR-2004 (first entry)

XX DE Single chain antibody that immunospecifically binds BlyS SeqID 1115.

XX antibody; B lymphocyte stimulator; BlyS; tumour necrosis factor;
 KW B cell proliferation; differentiation; scFv; myasthenia gravis;
 KW multiple sclerosis; asthma; rheumatoid arthritis; AIDS; leukaemia;
 KW carcinoma; lymphoma; antirheumatic; antiarthritic; neuroprotective;
 KW antiinflammatory; antiasthmatic; antiallergic; cytostatic.

XX OS Unidentified.

XX PN WO2003055979-A2.

XX PD 10-JUL-2003.

XX PF 14-NOV-2002; 2002WO-US036496.

XX PR 16-NOV-2001; 2001US-0331469P.

XX PR 19-DEC-2001; 2001US-0340817P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Ruben SM, Barash SC, Choi GH, Vaughan TU, Hilbert D;

XX DR WPI; 2003-505530/47.

XX Novel antibody that immunospecifically binds to a B lymphocyte stimulator
 PT (BlyS), useful for detecting and treating diseases or disorders e.g.
 PT rheumatoid arthritis, asthma and leukemia.

XX Example 1; SEQ ID NO 1115; 394pp; English.

CC This invention relates to novel antibodies that immunospecifically bind
 CC to B lymphocyte stimulator (BlyS). The BlyS gene has been mapped to
 CC chromosome 13q34 and encodes a protein that is a member of the tumour
 CC necrosis factor superfamily and induces both in vivo and in vitro B cell
 CC proliferation and differentiation. Specifically, it refers to single
 CC chain antibody molecules (scFvs) derived, preferably, from the variable
 CC heavy CDR3 region that immunospecifically bind to a polypeptide, or
 CC fragment thereof, of either human, murine, rat or monkey BlyS. The
 CC present invention refers to the use of such antibodies in various methods
 CC for the detection, diagnosis and prognosis of diseases related to the
 CC aberrant expression or inappropriate function of BlyS or its receptor. As
 CC such, these compositions are useful for identifying immune disorders
 CC including myasthenia gravis and multiple sclerosis, inflammatory
 CC disorders e.g. asthma and rheumatoid arthritis, infectious diseases such
 CC as AIDS and proliferative disorders including leukaemia, carcinoma and

CC lymphoma. Accordingly, they can be described as exhibiting various
 CC activities such as antirheumatic, antiarthritic, neuroprotective,
 CC antiinflammatory, antiasthmatic, antiallergic and cytostatic. This
 CC polypeptide sequence is a single chain antibody that binds Blys of the
 CC invention. NOTE: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 249 AA;

Query Match 70.8%; Score 46; DB 7; Length 249;
 Best Local Similarity 72.7%; Pred. No. 21;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NSWDSSGTHPV 11
 :|||||
 Db 228 HSWDSSGNHVV 238

RESULT 6
 ABP45105
 ID ABP45105 standard; protein; 247 AA.
 XX
 AC ABP45105;
 XX
 DT 19-AUG-2002 (first entry)
 XX
 DE Human Blys binding scFv SEQ ID 1116.
 XX
 KW Blys; B lymphocyte stimulator; TNF superfamily; human; cytostatic;
 KW tumour necrosis factor; B cell proliferation; B cell differentiation;
 KW immunosuppressive; immunostimulant; immunomodulatory; antirheumatic;
 KW antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;
 KW systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;
 KW common variable immunodeficiency; acquired immunodeficiency syndrome.
 XX
 OS Homo sapiens.
 XX
 FN WO200202641-A1.
 XX
 PD 10-JAN-2002.
 XX
 PF 15-JUN-2001; 2001WO-US019110.
 XX
 PR 16-JUN-2000; 2000US-0212210P.
 PR 17-OCT-2000; 2000US-0240816P.
 PR 16-MAR-2001; 2001US-0276249P.
 PR 21-MAR-2001; 2001US-0277379P.
 PR 25-MAY-2001; 2001US-0293499P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PA (CAMB-) CAMBRIDGE ANTIBODY TECHNOLOGY.
 XX
 PI Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;
 PI WPI; 2002-114799/15.
 XX
 DR Antibodies against B Lymphocyte Stimulating polypeptides, useful for the
 XX diagnosis and treatment of cancers and immune disorders.
 PT
 PS Claim 1; Page 1734-1735; 3148pp; English.
 XX
 CC This invention describes novel antibodies that immunospecifically bind to
 CC B Lymphocyte Stimulator (Blys) polypeptides. Blys is a member of the
 CC tumour necrosis factor (TNF) super family and induces B cell
 CC proliferation and differentiation. The antibodies of the invention have
 CC cytostatic, immunosuppressive, immunostimulant, immunomodulatory,
 CC antirheumatic and antiAIDS activity and can be used in vaccines to
 CC inhibit the expression and activity of Blys. The antibodies bind to Blys
 CC and so may be used to detect and quantitate the presence of Blys in
 CC biological samples and may be used in this way to diagnose disease
 CC associated with aberrant expression of Blys. They may also be
 CC administered to treat diseases associated with aberrant Blys expression

CC and activity such as cancer, immune, and autoimmune disorders and
 CC diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis,
 CC immunodeficiency (e.g. common variable immunodeficiency (CVID) and
 CC acquired immunodeficiency syndrome (AIDS)). ABP43990-ABP47228 represent
 CC the antibodies and fragments of the antibodies described in the method of
 CC the invention
 XX
 SQ Sequence 247 AA;

Query Match 67.7%; Score 44; DB 5; Length 247;
 Best Local Similarity 77.8%; Pred. No. 44;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NSWDSSGTH 9
 :|||||
 Db 226 HSWDSSGNH 234

RESULT 7
 ADG95932
 ID ADG95932 standard; protein; 247 AA.
 XX
 AC ADG95932;
 XX
 DT 11-MAR-2004 (first entry)
 XX
 DE Single chain antibody that immunospecifically binds Blys SeqID 1116.
 XX
 KW antibody; B lymphocyte stimulator; Blys; tumour necrosis factor;
 KW B cell proliferation; differentiation; scFv; myasthenia gravis;
 KW multiple sclerosis; asthma; rheumatoid arthritis; AIDS; leukaemia;
 KW carcinoma; lymphoma; antirheumatic; antiallergic; neuroprotective;
 KW antinflammatory; antiasthmatic; antiallergic; cycostatic.
 XX
 OS Unidentified.
 XX
 FN WO2003055979-A2.
 XX
 PD 10-JUL-2003.
 XX
 PF 14-NOV-2002; 2002WO-US036496.
 XX
 PR 16-NOV-2001; 2001US-0331469P.
 PR 19-DEC-2001; 2001US-0340817P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Ruben SM, Barash SC, Choi GH, Vaughan TJ, Hilbert D;
 XX WPI; 2003-505530/47.
 DR
 XX
 PT Novel antibody that immunospecifically binds to a B lymphocyte stimulator
 PT (Blys), useful for detecting and treating diseases or disorders e.g.
 PT rheumatoid arthritis, asthma and leukemia.
 PT
 PS Example 1; SEQ ID NO 1116; 394pp; English.
 XX
 CC This invention relates to novel antibodies that immunospecifically bind
 CC to B lymphocyte stimulator (Blys). The Blys gene has been mapped to
 CC chromosome 13q34 and encodes a protein that is a member of the tumour
 CC necrosis factor superfamily and induces both in vivo and in vitro B cell
 CC proliferation and differentiation. Specifically, it refers to single
 CC chain antibody molecules (scFvs) derived, preferably, from the variable
 CC heavy CDR3 region that immunospecifically bind to a polypeptide, or
 CC fragment thereof, of either human, murine, rat or monkey Blys. The
 CC present invention refers to the use of such antibodies in various methods
 CC for the detection, diagnosis and prognosis of diseases related to the
 CC aberrant expression or inappropriate function of Blys or its receptor. As
 CC such, these compositions are useful for identifying immune disorders
 CC including myasthenia gravis and multiple sclerosis, inflammatory
 CC disorders e.g. asthma and rheumatoid arthritis, infectious diseases such
 CC as AIDS and proliferative disorders including leukaemia, carcinoma and
 CC lymphoma. Accordingly, they can be described as exhibiting various

CC activities such as antirheumatic, antiarthritic, neuroprotective,
 CC antiinflammatory, antiaesthetic, antiallergic and cytostatic. This
 CC polypeptide sequence is a single chain antibody that binds Blys of the
 CC invention. NOTE: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 247 AA;

Query Match 67.7%; Score 44; DB 7; Length 247;
 Best Local Similarity 77.8%; Pred. No. 44;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NSWSSSGTH 9

Db 226 HSWSSGNH 234

RESULT 8

ADSO9248
 ID ADS09248 standard; protein; 245 AA.

XX AC ADS09248;

DT 18-NOV-2004 (first entry)

DE Human c-Met protein tyrosine kinase antibody, PGIA-1-A10.

KW c-Met; tyrosine kinase antibody; antigen binding; cytostatic;
 KW ophthalmological; antiinflammatory; analgesic; vasotropic; antipsoriatic;
 KW osteopathic; cancer; tumor; ophthalmic disease; glaucoma; retinitis;
 KW retinopathy; uveitis; ocular photophobia; macular degeneration; pain;
 KW acute injury; eye; hyperproliferative disorder; restenosis; angioplasty;
 KW psoriasis; HGF; osteoporosis; cancer.

XX OS Homo sapiens.

XX FN WO2004072117-A2.

XX PD 26-AUG-2004.

XX PF 11-FEB-2004; 2004WO-IB000503.

XX PR 13-FEB-2003; 2003US-0447073P.

XX PA (PHAA) PHARMACIA CORP.

XX PI Morton PA, Arbuckle JA, Evans ML, Joy WD, Kahn LE, Shieh JJ;

XX DR WPI; 2004-616044/59.

XX DR N-PSDB; ADS09308.

XX PT Novel c-Met protein tyrosine kinase antibody or its antigen-binding
 PT portion specifically binding to c-Met, useful for manufacture of
 PT medicament for treating cancer or tumor and for treatment of ophthalmic
 PT diseases such as glaucoma.

XX PS Claim 1; SEQ ID NO 10; 303pp; English.

XX The invention relates to a novel c-Met protein tyrosine kinase antibody
 CC or its antigen binding portion that specifically binds to c-Met. The c-
 CC Met antibody comprises any one of 1-60 fully defined sequence of 238,
 CC 244, 240, 250, 251, 242, 245, 247, 246, 253, 249, 243, 241, etc., amino
 CC acids as given in the specification, or its fragment. The invention
 CC further comprises: a pharmaceutical composition comprising the c-Met
 CC protein tyrosine kinase antibody and a carrier; an isolated cell that
 CC produces the c-Met protein tyrosine kinase antibody; and an isolated
 CC nucleic acid molecule that comprises a nucleic acid sequence that encodes
 CC a heavy chain or its antigen-binding portion or light chain or its
 CC antigen-binding portion of the c-Met protein tyrosine kinase antibody.
 CC The c-Met protein tyrosine kinase antibody has cytostatic,
 CC ophthalmological, antiinflammatory, analgesic, vasotropic, antipsoriatic,
 CC and osteopathic activities. The c-Met protein tyrosine kinase antibody is

CC useful for the manufacture of medicament for the treatment of cancer or
 CC tumor. The c-Met protein tyrosine kinase antibody is useful for
 CC diagnosing the presence or ligation of c-Met expressing tissue. The c-Met
 CC protein tyrosine kinase antibody is useful for detecting c-Met in a
 CC biological sample in vitro or in vivo. The c-Met protein tyrosine kinase
 CC antibody is also useful in the treatment or prevention of ophthalmic
 CC diseases such as glaucoma, retinitis, retinopathies (e.g., diabetic
 CC retinopathy), uveitis, ocular photophobia, macular degeneration and pain
 CC associated with acute injury to the eye. The pharmaceutical composition
 CC is useful for the treatment of hyperproliferative disorders such as
 CC restenosis after angioplasty, and psoriasis, and for the treatment of
 CC animals that lack sufficient HGF, e.g. osteoporosis and cancer. This
 CC sequence represents the protein of a phage display generated human c-Met
 CC antibody of the invention.

XX SQ Sequence 245 AA;

Query Match 66.2%; Score 43; DB 8; Length 245;
 Best Local Similarity 77.8%; Pred. NO. 65;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 WSSGTHPV 11

Db 226 WSSSDHPV 234

RESULT 9

ABB06006
 ID ABB06006 standard; protein; 508 AA.

XX AC ABB06006;

XX DT 09-MAY-2002 (first entry)

XX DE Human cancer suppressing protein PP844 SEQ ID NO:26.

XX KW Human; cancer suppression; cancer.

XX OS Homo sapiens.

XX FN CN1313318-A.

XX PD 19-SEP-2001.

XX PF 14-MAR-2000; 2000CN-00111997.

XX PR 14-MAR-2000; 2000CN-00111997.

XX PA (SHAN-) SHANGHAI INST ONCOLOGY.

XX PI Gu J, Yang S;

XX DR WPI; 2002-042196/06.

XX DR N-PSDB; ABL39592.

XX PT New human protein able to suppress growth of cancer cells and its
 PT encoding polynucleotide.

XX PS Claim 1; Page 39 (Disclosure); 65pp; Chinese.

XX The present invention describes human proteins with cancer suppressing
 CC activity. Also described are the polynucleotides encoding the proteins
 CC and a process for preparing the proteins by DNA recombination. The
 CC proteins and polynucleotides can be used in the treatment of diseases
 CC such as cancer. The present sequence represents a human cancer
 CC suppressing protein from the present invention

XX SQ Sequence 508 AA;

Query Match 66.2%; Score 43; DB 5; Length 508;
 Best Local Similarity 70.0%; Pred. No. 1.4e+02;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 NSWDSSGTHP 10
 ||| ||| |||
 Db 160 NSWDSSSRHP 169

RESULT 10
 ABP45969
 ID ABP45969 standard; protein; 246 AA.
 XX
 AC ABP45969;
 XX
 DT 19-AUG-2002 (first entry)
 XX
 DE Human BlyS binding scFv SEQ ID 1980.
 XX
 KW BlyS; B lymphocyte stimulator; TNF superfamily; human; cytostatic;
 KW tumour necrosis factor; B cell proliferation; B cell differentiation;
 KW immunosuppressive; immunostimulant; immunomodulatory; antirheumatic;
 KW antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;
 KW systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;
 KW common variable immunodeficiency; acquired immunodeficiency syndrome.
 XX
 OS Homo sapiens.
 XX
 PN WO200202641-A1.
 XX
 PD 10-JAN-2002.
 XX
 PF 15-JUN-2001; 2001WO-US019110.
 XX
 PR 16-JUN-2000; 2000US-0212210P.
 PR 17-OCT-2000; 2000US-0240816P.
 PR 16-MAR-2001; 2001US-0276248P.
 PR 21-MAR-2001; 2001US-0277379P.
 PR 25-MAY-2001; 2001US-0293499P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PA (CAMB-) CAMBRIDGE ANTIBODY TECHNOLOGY.
 XX
 PI Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;
 XX
 DR WPI; 2002-114799/15.
 XX
 PT Antibodies against B Lymphocyte Stimulating polypeptides, useful for the
 PT diagnosis and treatment of cancers and immune disorders.
 PS
 PS Claim 1; Page 2763-2764; 3148pp; English.
 XX
 CC This invention describes novel antibodies that immunospecifically bind to
 CC B Lymphocyte Stimulator (BlyS) polypeptides. BlyS is a member of the
 CC tumour necrosis factor (TNF) super family and induces B cell
 CC proliferation and differentiation. The antibodies of the invention have
 CC cytostatic, immunosuppressive, immunostimulant, immunomodulatory,
 CC antirheumatic and antiAIDS activity and can be used in vaccines to
 CC inhibit the expression and activity of BlyS. The antibodies bind to BlyS
 CC and so may be used to detect and quantitate the presence of BlyS in
 CC biological samples and may be used in this way to diagnose disease
 CC associated with aberrant expression of BlyS. They may also be
 CC administered to treat diseases associated with aberrant BlyS expression
 CC and activity such as cancer, immune, and autoimmune disorders and
 CC diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis,
 CC immunodeficiency (e.g. common variable immunodeficiency (CVID) and
 CC acquired immunodeficiency syndrome (AIDS)). ABP43990-ABP47228 represent
 CC the antibodies and fragments of the antibodies described in the method of
 CC the invention
 XX
 SQ Sequence 246 AA;

Query Match 63.1%; Score 41; DB 5; Length 246;
 Best Local Similarity 81.8%; Pred. No. 1.4e+02;
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NSWDSSGTHPV 11

Db 225 NSRDSSGTHLV 235
 ||| ||| ||| |||
 RESULT 11
 ADG96796
 ID ADG96796 standard; protein; 246 AA.
 XX
 AC ADG96796;
 XX
 DT 11-MAR-2004 (first entry)
 XX
 DE Single chain antibody that immunospecifically binds BlyS SeqID 1980.
 XX
 KW antibody; B lymphocyte stimulator; BlyS; tumour necrosis factor;
 KW B cell proliferation; differentiation; scFv; myasthenia gravis;
 KW multiple sclerosis; asthma; rheumatoid arthritis; AIDS; leukaemia;
 KW carcinoma; lymphoma; antirheumatic; antiallergic; neuroprotective;
 KW antiinflammatory; antiasthmatic; antiallergic; cytostatic.
 XX
 OS Unidentified.
 XX
 PN WO2003055979-A2.
 XX
 PD 10-JUL-2003.
 XX
 PF 14-NOV-2002; 2002WO-US036496.
 XX
 PR 16-NOV-2001; 2001US-0331469P.
 PR 19-DEC-2001; 2001US-0340817P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Ruben SM, Barash SC, Choi GH, Vaughan TJ, Hilbert D;
 XX
 DR WPI; 2003-505530/47.
 XX
 PT Novel antibody that immunospecifically binds to a B lymphocyte stimulator
 PT (BlyS), useful for detecting and treating diseases or disorders e.g.
 PT rheumatoid arthritis, asthma and leukemia.
 XX
 PS Example 1; SEQ ID NO 1980; 394pp; English.
 XX
 CC This invention relates to novel antibodies that immunospecifically bind
 CC to B lymphocyte stimulator (BlyS). The BlyS gene has been mapped to
 CC chromosome 13q34 and encodes a protein that is a member of the tumour
 CC necrosis factor superfamily and induces both in vivo and in vitro B cell
 CC proliferation and differentiation. Specifically, it refers to single
 CC chain antibody molecules (scFvs) derived, preferably, from the variable
 CC heavy CDR3 region that immunospecifically bind to a polypeptide, or
 CC fragment thereof, of either human, murine, rat or monkey BlyS. The
 CC present invention refers to the use of such antibodies in various methods
 CC for the detection, diagnosis and prognosis of diseases related to the
 CC aberrant expression or inappropriate function of BlyS or its receptor. As
 CC such, these compositions are useful for identifying immune disorders
 CC including myasthenia gravis and multiple sclerosis, inflammatory
 CC disorders e.g. asthma and rheumatoid arthritis, infectious diseases such
 CC as AIDS and proliferative disorders including leukaemia, carcinoma and
 CC lymphoma. Accordingly, they can be described as exhibiting various
 CC activities such as antirheumatic, antiallergic, neuroprotective,
 CC antiinflammatory, antiasthmatic, antiallergic and cytostatic. This
 CC polypeptide sequence is a single chain antibody that binds BlyS of the
 CC invention. NOTE: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published pct_sequences.
 XX
 SQ Sequence 246 AA;

Query Match 63.1%; Score 41; DB 7; Length 246;
 Best Local Similarity 81.8%; Pred. No. 1.4e+02;
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NSWDSSGTHPV 11


```

AAW15538
ID AAW15538 standard; protein; 110 AA.
XX AC AAW15538;
XX DT 27-NOV-1997 (first entry)
XX DE Anti-TGF beta-1 scFv antibody 10A6 VL domain.
XX KW Transforming growth factor beta-1; TGF-beta-1; human;
XX KW antibody engineering; scFv; phage display; lung fibrosis;
XX KW arterial injury; proliferative retinopathy; retinal detachment;
XX KW adult respiratory distress syndrome; liver cirrhosis;
XX KW post myocardial infarction; post-angioplasty restenosis; scleroderma;
XX KW vascular disease; cataract; glaucoma; scarring; glomerulonephritis;
XX KW osteoporosis; immune disease; inflammation; rheumatoid arthritis;
XX KW macrophage deficiency disease; macrophage pathogen infection; therapy.
XX OS Homo sapiens.
XX PN GB2305921-A.
XX PD 23-APR-1997.
XX PF 07-OCT-1996; 96GB-00020920.
XX PR 06-OCT-1995; 95GB-00020486.
XX PR 19-JAN-1996; 96GB-00001081.
XX PA (CAMB-) CAMBRIDGE ANTIBODY TECHNOLOGY.
XX PI Thompson JE, Vaughan TJ, Williams AJ, Green JA, Jackson RH;
XX PI Bacon L, Johnson KS, Wilton AJ, Tempest PR, Pope AR;
XX DR WPI; 1997-215360/20.
XX DR N-PSDB; AAT60384.
XX PT Agent contg. antigen-binding domain of human antibody to transforming
XX PT growth factor beta 1 or 2 - and nucleic acid encoding it, used to
XX PT neutralise effects of TGF, e.g. for control of fibrosis, immune and
XX PT inflammatory disease.
XX PS Claim 16; Fig 1b(ii); 184pp; English.
XX CC This polypeptide sequence comprises the VL domain of human scFv antibody
XX CC 10A6, which is specific for transforming growth factor (TGF) beta-1. It
XX CC is encoded by a gene (AAT60384) isolated from a peripheral blood
XX CC lymphocyte library. The antigen-binding domains of human antibodies (see
XX CC AAW15522-40) to TGF beta-1 and/or beta-2 can be used to counter the
XX CC adverse effects of TGF beta, such as (i) promotion of fibrosis (in
XX CC dermal, ocular or keloid scarring, lung fibrosis, arterial injury,
XX CC proliferative retinopathy, retinal detachment, adult respiratory distress
XX CC syndrome, liver cirrhosis, post myocardial infarction, post-angioplasty
XX CC restenosis, scleroderma, vascular disorders, cataract, glaucoma, or esp.
XX CC neural scarring and glomerulonephritis, also (not claimed) osteoporosis),
XX CC or (ii) immune and inflammatory diseases (e.g. rheumatoid arthritis,
XX CC macrophage deficiency diseases or macrophage pathogen infection). Nucleic
XX CC acids encoding human antibody VH and VL can be used for prodn. of
XX CC recombinant antigen-binding domains. These are highly specific, have low
XX CC dissociation constants (pref. less than 5 nM) and low IC50s for
XX CC neutralisation
XX CC Sequence 110 AA;
XX CC Query Match 61.5%; Score 40; DB 2; Length 110;
XX CC Best Local Similarity 88.9%; Pred. No. 85;
XX CC Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 NSWDSSGTH 9
DB 88 NSRDSGTH 96

AAW15538
ID AAW15538 standard; protein; 111 AA.
XX AC AAR12264;
XX DT 25-MAR-2003 (revised)
XX DT 15-AUG-1991 (first entry)
XX DE Anti-human RHD PAG-1 MAb (VL chain).
XX KW Monoclonal antibody; rhesus D; blood-typing; CDR;
XX KW haemolytic disease of the newborn; HDN.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT 23..33
XX FT /label= CDR1
XX FT 49..55
XX FT /label= CDR2
XX FT 88..100
XX FT /label= CDR3
XX PN W09107492-A.
XX PD 30-MAY-1991.
XX PF 13-NOV-1989; 89GB-00025590.
XX PR 13-NOV-1989; 89GB-00025590.
XX PA (BLOO-) CENT BLOOD LAB AUTHORITY.
XX PI Hughesjone N;
XX DR WPI; 1991-178104/24.
XX DR N-PSDB; AAQ11946.
XX PT DNA encoding complementary determining regions - of human anti-rhesus d
XX PT antibodies, useful in prodn. of monoclonal antibodies and for passive
XX PT immunisation.
XX PS Disclosure; Fig 3; 32pp; English.
XX CC The DNA sequence of eleven monoclonal antibodies are represented in
XX CC Q119145-57. Synthetic genes, for both heavy and light chains may be
XX CC created by combining selected CDR 1, 2, and 3 regions, which may be
XX CC selected from different antibody mols. having varied binding specificity.
XX CC The chimaeric anti-RHD antibodies can be used for diagnosis and therapy,
XX CC and are capable of providing blood- typing reagents of high specificity
XX CC and reliability. They can also be used in passive immunisation to prevent
XX CC haemolytic disease of the newborn. (Updated on 25-MAR-2003 to correct PA
XX CC field.) (Updated on 25-MAR-2003 to correct PI field.)
XX CC Sequence 111 AA;
XX CC Query Match 61.5%; Score 40; DB 2; Length 111;
XX CC Best Local Similarity 75.0%; Pred. No. 86;
XX CC Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 WDSGTHP 10
DB 90 WDSSAHP 97

Search completed: April 28, 2005, 18:17:24
Job time : 85.8438 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2005, 18:06:50 ; Search time 22.3438 Seconds
(without alignments)
36.750 Million cell updates/sec

Title: US-10-088-639A-2_COPY_88_98
Perfect score: 65
Sequence: 1 NSWSSGTHPV 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	61.5	13	4	US-09-563-222C-25
2	40	61.5	448	4	US-09-252-991A-24066
3	38	58.5	92	2	US-08-447-464-6
4	38	58.5	92	2	US-08-716-679-6
5	38	58.5	345	4	US-09-248-796A-20780
6	38	58.5	461	1	US-08-672-571A-3
7	38	58.5	490	1	US-08-672-571A-1
8	38	58.5	1075	4	US-09-949-016-8308
9	38	58.5	1452	2	US-08-652-971-4
10	38	58.5	1452	2	US-08-449-644-8
11	38	58.5	1452	2	US-08-087-244A-8
12	38	58.5	1452	2	US-08-991-258A-4
13	38	58.5	1452	2	US-08-769-339-4
14	38	58.5	1452	3	US-08-991-953A-4
15	37.5	57.7	749	4	US-09-562-737-97
16	37	56.9	11	2	US-08-350-260A-349
17	37	56.9	11	2	US-08-350-260A-430
18	37	56.9	11	4	US-09-104-337A-349
19	37	56.9	11	4	US-09-104-337A-430
20	37	56.9	11	4	US-09-205-258-1061
21	37	56.9	32	3	US-09-167-681-34
22	37	56.9	32	3	US-09-167-681-42
23	37	56.9	32	3	US-09-167-681-50
24	37	56.9	67	2	US-08-273-146-67
25	37	56.9	95	4	US-09-621-976-7103
26	37	56.9	97	2	US-08-665-202-35
27	37	56.9	97	4	US-09-315-574-35

28	37	56.9	101	4	US-09-726-219A-168	Sequence 168, App
29	37	56.9	103	4	US-09-673-395A-348	Sequence 348, App
30	37	56.9	108	4	US-09-720-493-4	Sequence 4, Appl
31	37	56.9	109	2	US-08-652-816A-16	Sequence 16, Appl
32	37	56.9	118	4	US-09-205-258-1060	Sequence 1060, Ap
33	37	56.9	124	4	US-03-231-077D-10	Sequence 10, Appl
34	37	56.9	242	4	US-09-543-681A-6843	Sequence 6843, Ap
35	37	56.9	270	4	US-09-809-920-14	Sequence 14, Appl
36	37	56.9	278	3	US-09-260-527-3	Sequence 3, Appl
37	37	56.9	280	3	US-09-260-527-1	Sequence 1, Appl
38	37	56.9	295	4	US-09-328-174A-3	Sequence 3, Appl
39	37	56.9	309	3	US-09-079-029-9	Sequence 9, Appl
40	37	56.9	312	3	US-09-079-029-10	Sequence 10, Appl
41	37	56.9	336	4	US-09-673-395A-600	Sequence 600, App
42	37	56.9	352	4	US-09-910-050B-18	Sequence 18, Appl
43	37	56.9	352	4	US-10-403-192-18	Sequence 18, Appl
44	37	56.9	356	4	US-09-252-991A-32092	Sequence 32092, A
45	37	56.9	403	4	US-09-252-991A-22238	Sequence 22238, A

ALIGNMENTS

RESULT 1

US-09-563-222C-25
; Sequence 25, Application US/09563222C
; Patent No. 6696620
; GENERAL INFORMATION:
; APPLICANT: EPICYTE PHARMACEUTICALS, INC.
; APPLICANT: HIAIT, ANDREW C.
; APPLICANT: HEIN, MICH B.
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN ARRAYS IN PLANT CELLS
; FILE REFERENCE: 068904-0501
; CURRENT APPLICATION NUMBER: US/09/563.222C
; CURRENT FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: PCT/US01/14349
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: 09/563,222
; PRIOR FILING DATE: 2000-05-02
; NUMBER OF SEQ ID NOS: 182
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-563-222C-25

Query Match 61.5%; Score 40; DB 4; Length 13;
Best Local Similarity 66.7%; Pred. No. 2.2;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 SWSSSGTHP 10
:|||||
Db 2 AWDSSSDHP 10

RESULT 2

US-09-252-991A-24066
; Sequence 24066, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24066

APPLICANT: Keith Weinstock et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICA
TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS

Query Match 58.5%; Score 38; DB 4; Length 345;
 Best Local Similarity 62.5%; Pred. NO. 1.5e+02;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 WDSGTHP 10
 Db 66 WDTGKHP 73

RESULT 6

US-08-672-571A-3
 ; Sequence 3, Application US/08672571A
 ; Patent No. 5795765
 ; GENERAL INFORMATION:
 ; APPLICANT: IZU, Hiroyuki
 ; APPLICANT: KURUME, Yoko
 ; APPLICANT: IZUMI, Yoshiya
 ; APPLICANT: SANO, Mutsumi
 ; APPLICANT: KATO, Ikunoshin
 ; TITLE OF INVENTION: Gene Encoding Endoglycoceramidase
 ; NUMBER OF SEQUENCES: 15
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
 ; STREET: P.O. Box 747
 ; CITY: Falls Church
 ; STATE: Virginia
 ; COUNTRY: USA
 ; ZIP: 22040-0747
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
 ; CURRENT APPLICATION DATA: US/08/672,571A
 ; FILING DATE: 28 JUNE 1996
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: WEINER, Marc S.
 ; REGISTRATION NUMBER: 32,181
 ; REFERENCE/DOCKET NUMBER: 1422-0264P
 ; TELEPHONE: (703) 205-8000
 ; TELEFAX: (703) 205-8050
 ; TELEX: 248345
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 461 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; US-08-672-571A-3

Query Match 58.5%; Score 38; DB 1; Length 461;
 Best Local Similarity 50.0%; Pred. NO. 2e+02;
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NSWDSSGTHP 10
 Db 164 NFWNTGKHP 173

RESULT 7

US-08-672-571A-1
 ; Sequence 1, Application US/08672571A
 ; Patent No. 5795765
 ; GENERAL INFORMATION:
 ; APPLICANT: IZU, Hiroyuki
 ; APPLICANT: KURUME, Yoko
 ; APPLICANT: IZUMI, Yoshiya
 ; APPLICANT: SANO, Mutsumi

; APPLICANT: KATO, Ikunoshin
 ; APPLICANT: ITO, Makoto
 ; TITLE OF INVENTION: Gene Encoding Endoglycoceramidase
 ; NUMBER OF SEQUENCES: 15
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
 ; STREET: P.O. Box 747
 ; CITY: Falls Church
 ; STATE: Virginia
 ; COUNTRY: USA
 ; ZIP: 22040-0747
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
 ; CURRENT APPLICATION DATA: US/08/672,571A
 ; FILING DATE: 28 JUNE 1996
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: WEINER, Marc S.
 ; REGISTRATION NUMBER: 32,181
 ; REFERENCE/DOCKET NUMBER: 1422-0264P
 ; TELEPHONE: (703) 205-8000
 ; TELEFAX: (703) 205-8050
 ; TELEX: 248345
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 490 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; US-08-672-571A-1

Query Match 58.5%; Score 38; DB 1; Length 490;
 Best Local Similarity 50.0%; Pred. NO. 2.1e+02;
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NSWDSSGTHP 10
 Db 193 NFWNTGKHP 202

RESULT 8

US-09-949-016-8308
 ; Sequence 8308, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 8308
 ; LENGTH: 1075
 ; TYPE: PRT
 ; ORGANISM: Human
 ; US-09-949-016-8308

Query Match 58.5%; Score 38; DB 4; Length 1075;
 Best Local Similarity 55.6%; Pred. NO. 4.8e+02;

Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 SWDSSGTHP 10
|||: :||
Db 57 SWDTNSHP 65

RESULT 9
US-08-652-971-4
; Sequence 4, Application US/08652971
; Patent No. 5814507
; GENERAL INFORMATION:
; APPLICANT: Cheng, Jill
; APPLICANT: Lasky, Laurence A.
; TITLE OF INVENTION: A NOVEL KAPPA/MU-LIKE PROTEIN TYROSINE
; TITLE OF INVENTION: PHOSPHATASE, PTP LAMBDA
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd.
; CITY: South San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/652,971
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Dregger, Ginger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: P1033
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 225-3216
; TELEFAX: (415) 952-9881
; TELEX: 910 371-7168
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1452 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-652-971-4

Query Match 58.5%; Score 38; DB 2; Length 1452;
Best Local Similarity 55.6%; Pred. No. 6.6e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 SWDSSGTHP 10
|||: :||
Db 434 SWDTNSHP 442

RESULT 10
US-08-449-644-8
; Sequence 8, Application US/08449644
; Patent No. 5856162
; GENERAL INFORMATION:
; APPLICANT: Schllessinger, Joseph
; APPLICANT: Sap, Jan M.
; APPLICANT: Ullrich, Axel
; APPLICANT: Vogel, Wolfgang
; APPLICANT: Fuchs, Miriam
; TITLE OF INVENTION: NOVEL RECEPTOR-TYPE PHOSPHOTYROSINE
; TITLE OF INVENTION: PHOSPHATASE-KAPPA
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:

ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/449,644
; FILING DATE: 24-MAY-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/087,244
; FILING DATE: 01-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7683-042
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1452 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-449-644-8

Query Match 58.5%; Score 38; DB 2; Length 1452;
Best Local Similarity 55.6%; Pred. No. 6.6e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 SWDSSGTHP 10
|||: :||
Db 434 SWDTNSHP 442

RESULT 11
US-08-087-244A-8
; Sequence 8, Application US/08087244A
; Patent No. 5863755
; GENERAL INFORMATION:
; APPLICANT: Schllessinger, Joseph
; APPLICANT: Sap, Jan M.
; APPLICANT: Ullrich, Axel
; APPLICANT: Vogel, Wolfgang
; APPLICANT: Fuchs, Miriam
; TITLE OF INVENTION: NOVEL RECEPTOR-TYPE PHOSPHOTYROSINE
; TITLE OF INVENTION: PHOSPHATASE-KAPPA
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/087,244A
; FILING DATE: 01-JUL-1993
; CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7683-042
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-790-9090
TELEFAX: 212-869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1452 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-087-244A-8

Query Match 58.5%; Score 38; DB 2; Length 1452;
Best Local Similarity 55.6%; Pred. No. 6.6e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 SWDSSGTHP 10
Db 434 SWDTNSHP 442

RESULT 12
US-08-991-258A-4
Sequence 4, Application US/08991258A
Patent No. 5928887
GENERAL INFORMATION:
APPLICANT: Cheng, Jill
APPLICANT: Lasky, Laurence A.
TITLE OF INVENTION: A NOVEL KAPPA/MU-LIKE PROTEIN TYROSINE
TITLE OF INVENTION: PHOSPHATASE, PTP LAMBDA
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT, LLP
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/991.258A
FILING DATE: 17-DEC-1997

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/652,971
FILING DATE: 24-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Dreger, Walter H.
REGISTRATION NUMBER: 24,190
REFERENCE/DOCKET NUMBER: A-63478-3/WHF/MTK
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX:

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1452 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-991-258A-4

Query Match 58.5%; Score 38; DB 2; Length 1452;

Best Local Similarity 55.6%; Pred. No. 6.6e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Qy 2 SWDSSGTHP 10
Db 434 SWDTNSHP 442

RESULT 13
US-08-769-399-4
Sequence 4, Application US/08769399
Patent No. 5976852
GENERAL INFORMATION:
APPLICANT: Cheng, Jill
APPLICANT: Lasky, Laurence A.
TITLE OF INVENTION: A NOVEL KAPPA/MU-LIKE PROTEIN TYROSINE
TITLE OF INVENTION: PHOSPHATASE, PTP LAMBDA
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd.
CITY: South San Francisco
STATE: California
COUNTRY: United States
ZIP: 94080

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/769.399
FILING DATE:

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Dreger, Ginger R.
REGISTRATION NUMBER: 33,055
REFERENCE/DOCKET NUMBER: P1033
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 225-3216
TELEFAX: (415) 952-9881
TELEX: 910 371-7168
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1452 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-769-399-4

Query Match 58.5%; Score 38; DB 2; Length 1452;
Best Local Similarity 55.6%; Pred. No. 6.6e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 SWDSSGTHP 10
Db 434 SWDTNSHP 442

RESULT 14
US-08-991-953A-4
Sequence 4, Application US/08991953A
Patent No. 6083748
GENERAL INFORMATION:
APPLICANT: Cheng, Jill
APPLICANT: Lasky, Laurence A.
TITLE OF INVENTION: A NOVEL KAPPA/MU-LIKE PROTEIN TYROSINE
TITLE OF INVENTION: PHOSPHATASE, PTP LAMBDA
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT, LLP
STREET: 4 Embarcadero Center, Suite 3400

Search completed: April 28, 2005, 18:28:06
Job time : 23:3438 secs

CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/991,953A
FILING DATE: 16-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/652,971
FILING DATE: 24-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Dreger, Walter H.
REGISTRATION NUMBER: 24,190
REFERENCE/POCKET NUMBER: A-63478-3/WH/MTK
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1452 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-991-953A-4

Query Match 58.5%; Score 38; DB 3; Length 1452;
Best Local Similarity 55.6%; Pred. No. 6.6e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 SWDSGTHP 10
Db 434 SWTDNSHP 442

RESULT 15
US-09-562-737-97
; Sequence 97, Application US/09562737
; Patent No. 6428967
; GENERAL INFORMATION:
; APPLICANT: Herz, Joachim
; APPLICANT: Gotthardt, Michael
; TITLE OF INVENTION: LDL Receptor Signaling Pathways
; FILE REFERENCE: UTSW0708
; CURRENT APPLICATION NUMBER: US/09/562,737
; CURRENT FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 97
; LENGTH: 749
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Sequence
US-09-562-737-97

Query Match 57.7%; Score 37.5; DB 4; Length 749;
Best Local Similarity 77.8%; Pred. No. 4e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

Qy 3 W-DSSGTHP 10
Db 141 WTDNSAGTHP 149

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2005, 18:24:46 ; Search time 68.2344 Seconds
(without alignments)
53.700 Million cell updates/sec

Title: US-10-088-639A-2_COPY_88_98
Perfect score: 65
Sequence: 1 NSWDSSGTHPV 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1426032 seqs, 333106140 residues

Total number of hits satisfying chosen parameters: 1426032

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US05_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US04_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	51	78.5	11	9	US-09-817-661-21
2	49	75.4	110	15	US-10-251-085B-173
3	49	75.4	110	16	US-10-737-252-173
4	49	75.4	168	15	US-10-424-599-177279
5	46	70.8	249	10	US-09-880-748-1115
6	46	70.8	249	15	US-10-293-418-1115
7	44	67.7	247	10	US-09-880-748-1116
8	44	67.7	247	15	US-10-293-418-1116
9	43	66.2	245	16	US-10-779-461-10
10	41	63.1	183	16	US-10-437-963-194069
11	41	63.1	246	10	US-09-880-748-1980
12	41	63.1	246	15	US-10-293-418-1980
13	40	61.5	13	10	US-09-563-222-25

Sequence 25, Appl
Sequence 243244,
Sequence 18, Appl
Sequence 18, Appl
Sequence 16, Appl
Sequence 353, App
Sequence 540, App
Sequence 540, App
Sequence 28149, A
Sequence 273242,
Sequence 626, App
Sequence 626, App
Sequence 164780,
Sequence 3149, Ap
Sequence 2, Appli
Sequence 250, App
Sequence 250, App
Sequence 274871,
Sequence 50343, A
Sequence 166593,
Sequence 120488,
Sequence 53692, A
Sequence 188, App
Sequence 188, App
Sequence 69140, A
Sequence 8523, Ap
Sequence 6568, Ap
Sequence 22, Appl
Sequence 128, App
Sequence 576, App
Sequence 92, Appl

ALIGNMENTS

RESULT 1

US-09-817-661-21
; Sequence 21, Application US/09817661
; Patent No. US20020076692A1
; GENERAL INFORMATION:
; APPLICANT: Osbourn, Jane
; APPLICANT: Hollet, Thor
; TITLE OF INVENTION: Improvements to ribosome display
; FILE REFERENCE: 84633
; CURRENT APPLICATION NUMBER: US/09/817,661
; CURRENT FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: US 60/193,802
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 21
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-817-661-21

Query Match 78.5%; Score 51; DB 9; Length 11;
Best Local Similarity 81.8%; Pred. No. 0.17;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NSWDSSGTHPV 11
| | | | | | | | | |
Db 1 NSWDSSGNHVV 11

RESULT 2

US-10-251-085B-173
; Sequence 173, Application US/10251085B
; Publication No. US20040072164A1
; GENERAL INFORMATION:
; APPLICANT: Bowdish, Katherine S.

```

; APPLICANT: Frederickson, Shana
; APPLICANT: Renshaw, Mark
; APPLICANT: Lin, Ying-Chi
; APPLICANT: Maruyama, Toshiaki
; TITLE OF INVENTION: ENGINEERED TEMPLATES AND THEIR USE IN SINGLE PRIMER AMPLIFICATION
; FILE REFERENCE: 1087-21
; CURRENT APPLICATION NUMBER: US/10/251,085B
; CURRENT FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 60/323,455
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 278
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 173
; LENGTH: 110
; TYPE: PRT
; ORGANISM: human
US-10-251-085B-173

Query Match      75.4%; Score 49; DB 15; Length 110;
Best Local Similarity 88.9%; Pred. No. 3.3;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      3 WDSSGTHPV 11
Db      92 WDSSGDHVF 100

RESULT 3
US-10-737-252-173
; Sequence 173, Application US/10737252
; Publication No. US20040175736A1
; GENERAL INFORMATION:
; APPLICANT: Bowdish, Katherine S.
; APPLICANT: Frederickson, Shana
; APPLICANT: Renshaw, Mark
; APPLICANT: Lin, Ying-Chi
; APPLICANT: Maruyama, Toshiaki
; TITLE OF INVENTION: ENGINEERED TEMPLATES AND THEIR USE IN SINGLE PRIMER AMPLIFICATION
; FILE REFERENCE: 1087-21 CIP
; CURRENT APPLICATION NUMBER: US/10/737,252
; CURRENT FILING DATE: 2003-12-15
; PRIOR APPLICATION NUMBER: US 10/251,085
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 60/323,455
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 309
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 173
; LENGTH: 110
; TYPE: PRT
; ORGANISM: human
US-10-737-252-173

Query Match      75.4%; Score 49; DB 16; Length 110;
Best Local Similarity 88.9%; Pred. No. 3.3;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      3 WDSSGTHPV 11
Db      92 WDSSGDHVF 100

RESULT 4
US-10-424-599-177279
; Sequence 177279, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kowalc David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

```

```

; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 177279
; LENGTH: 168
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_13109C.1.pep
US-10-424-599-177279

Query Match      75.4%; Score 49; DB 15; Length 168;
Best Local Similarity 70.0%; Pred. No. 5;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy      1 NSWDSSGTHP 10
Db      154 NSWESQSHP 163

RESULT 5
US-09-880-748-1115
; Sequence 1115, Application US/09880748
; Publication No. US2003005937A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind BlyS
; FILE REFERENCE: PF523
; CURRENT APPLICATION NUMBER: US/09/880,748
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 3239
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1115
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-748-1115

Query Match      70.8%; Score 46; DB 10; Length 249;
Best Local Similarity 72.7%; Pred. No. 22;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy      1 NSWDSSGTHPV 11
Db      228 HSWDSSGNHVV 238

RESULT 6
US-10-293-418-1115
; Sequence 1115, Application US/10293418
; Publication No. US20030223996A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind BlyS
; FILE REFERENCE: PF523P2
; CURRENT APPLICATION NUMBER: US/10/293,418
; CURRENT FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: 60/331,469
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/340,817
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 09/880,748

```

; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1115
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-418-1115

Query Match 70.8%; Score 46; DB 15; Length 249;
Best Local Similarity 72.7%; Pred. No. 22;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 NSWSSGTHPV 11
:|||||
DB 228 HSWDSSGNHV 238

RESULT 7

US-09-880-748-1116
; Sequence 1116, Application US/09880748
; Publication No. US2003005937A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind BlyS
; FILE REFERENCE: PF523
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 3239
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1116
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-748-1116

Query Match 67.7%; Score 44; DB 10; Length 247;
Best Local Similarity 77.8%; Pred. No. 46;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NSWSSGTH 9
:|||||
DB 226 HSWDSSGNH 234

RESULT 8

US-10-293-418-1116
; Sequence 1116, Application US/10293418
; Publication No. US2003023996A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind BlyS
; FILE REFERENCE: PF523P2
; CURRENT APPLICATION NUMBER: US/10/293,418

; CURRENT FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: 60/331,469
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/340,817
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 09/880,748
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1116
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-418-1116

Query Match 67.7%; Score 44; DB 15; Length 247;
Best Local Similarity 77.8%; Pred. No. 46;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NSWSSGTH 9
:|||||
DB 226 HSWDSSGNH 234

RESULT 9

US-10-779-461-10
; Sequence 10, Application US/10779461
; Publication No. US2004016654A1
; GENERAL INFORMATION:
; APPLICANT: Morton, Philip A
; TITLE OF INVENTION: ANTIBODIES TO c-MET FOR THE TREATMENT OF CANCERS
; FILE REFERENCE: 00980/1
; CURRENT APPLICATION NUMBER: US/10/779,461
; CURRENT FILING DATE: 2004-02-13
; PRIOR APPLICATION NUMBER: 60/447,073
; PRIOR FILING DATE: 2003-02-13
; NUMBER OF SEQ ID NOS: 161
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 10
; LENGTH: 245
; TYPE: PRT
; ORGANISM: artificial
; FEATURE:
; OTHER INFORMATION: phage display generated human antibody
US-10-779-461-10

Query Match 66.2%; Score 43; DB 16; Length 245;
Best Local Similarity 77.8%; Pred. No. 65;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 WSSSGTHPV 11
:|||||
DB 226 WSSSDHVP 234

RESULT 10

US-10-437-963-194069
; Sequence 194069, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei
 ; APPLICANT: Boukharov, Andrey A.
 ; APPLICANT: Barbazuk, Brad
 ; APPLICANT: Li, Ping
 ; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53221)B
 ; CURRENT APPLICATION NUMBER: US/10/437,963
 ; CURRENT FILING DATE: 2003-05-14
 ; NUMBER OF SEQ ID NOS: 204966
 ; SEQ ID NO 194069
 ; LENGTH: 183
 ; TYPE: PRT
 ; ORGANISM: Oryza sativa
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT4530_90147C.1.pap
 US-10-437-963-194069

Query Match 63.1%; Score 41; DB 16; Length 183;
 Best Local Similarity 70.0%; Pred. No. 1e+02;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 NSWDSSGTHP 10
 |||||
 Db 49 NSWPESGTSP 58

RESULT 11
 US-03-880-748-1980
 ; Sequence 1980, Application US/09880748
 ; Publication No. US20030059937A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ruben et al.
 ; TITLE OF INVENTION: Antibodies that Immunosepecifically Bind BlyS

; FILE REFERENCE: PF523
 ; CURRENT APPLICATION NUMBER: US/09/880,748
 ; CURRENT FILING DATE: 2001-06-15
 ; PRIOR APPLICATION NUMBER: 60/212,210
 ; PRIOR FILING DATE: 2000-06-15
 ; PRIOR APPLICATION NUMBER: 60/240,816
 ; PRIOR FILING DATE: 2000-10-17
 ; PRIOR APPLICATION NUMBER: 60/276,248
 ; PRIOR FILING DATE: 2001-03-16
 ; PRIOR APPLICATION NUMBER: 60/277,379
 ; PRIOR FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: 60/293,499
 ; PRIOR FILING DATE: 2001-05-25
 ; NUMBER OF SEQ ID NOS: 3239
 ; SOFTWARE: Patentin Ver. 2.0
 ; SEQ ID NO 1980
 ; LENGTH: 246
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens

US-03-880-748-1980
 Query Match 63.1%; Score 41; DB 10; Length 246;
 Best Local Similarity 81.8%; Pred. No. 1.4e+02;
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NSWDSSGTHP 11
 |||||
 Db 225 NSRDSSGTHLV 235

RESULT 12
 US-10-293-418-1980
 ; Sequence 1980, Application US/10293418
 ; Publication No. US20030223996A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ruben et al.
 ; TITLE OF INVENTION: Antibodies that Immunosepecifically Bind BlyS
 ; FILE REFERENCE: PF523P2
 ; CURRENT APPLICATION NUMBER: US/10/293,418

; CURRENT FILING DATE: 2002-11-27
 ; PRIOR APPLICATION NUMBER: 60/331,469
 ; PRIOR FILING DATE: 2001-11-16
 ; PRIOR APPLICATION NUMBER: 60/340,817
 ; PRIOR FILING DATE: 2001-12-19
 ; PRIOR APPLICATION NUMBER: 09/880,748
 ; PRIOR FILING DATE: 2001-06-15
 ; PRIOR APPLICATION NUMBER: 60/293,499
 ; PRIOR FILING DATE: 2001-05-25
 ; PRIOR APPLICATION NUMBER: 60/277,379
 ; PRIOR FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: 60/276,248
 ; PRIOR FILING DATE: 2001-03-16
 ; PRIOR APPLICATION NUMBER: 60/240,816
 ; PRIOR FILING DATE: 2000-10-17
 ; PRIOR APPLICATION NUMBER: 60/212,210
 ; PRIOR FILING DATE: 2000-06-16
 ; NUMBER OF SEQ ID NOS: 3247
 ; SEQ ID NO 1980
 ; LENGTH: 246
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-293-418-1980

Query Match 63.1%; Score 41; DB 15; Length 246;
 Best Local Similarity 81.8%; Pred. No. 1.4e+02;
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NSWDSSGTHPV 11
 |||||
 Db 225 NSRDSSGTHLV 235

RESULT 13
 US-09-563-222-25
 ; Sequence 25, Application US/09563222
 ; Publication No. US20030079253A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hiatt, Andrew
 ; APPLICANT: Hein, Mich B.
 ; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN ARRAYS IN
 ; TITLE OF INVENTION: EUKARYOTIC CELLS
 ; FILE REFERENCE: 310098.406
 ; CURRENT APPLICATION NUMBER: US/09/563,222
 ; CURRENT FILING DATE: 2000-05-02
 ; NUMBER OF SEQ ID NOS: 197
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 25
 ; LENGTH: 13
 ; TYPE: PRT
 ; ORGANISM: Homo sapien
 US-09-563-222-25

Query Match 61.5%; Score 40; DB 10; Length 13;
 Best Local Similarity 66.7%; Pred. No. 11;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 SWDSSGTHP 10
 :|||
 Db 2 AWDSSSDHP 10

RESULT 14
 US-10-783-950-25
 ; Sequence 25, Application US/10783950
 ; Publication No. US20040199945A1
 ; GENERAL INFORMATION:
 ; APPLICANT: EPICYTE PHARMACEUTICALS, INC.
 ; APPLICANT: HIATT, ANDREW C.
 ; APPLICANT: HEIN, MICH B.
 ; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN ARRAYS IN PLANT CELLS
 ; FILE REFERENCE: 068904-0501
 ; CURRENT APPLICATION NUMBER: US/10/783,950

; CURRENT FILING DATE: 2004-02-19
; PRIOR APPLICATION NUMBER: US/09/563,222
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: PCT/US01/14349
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: 09/563,222
; PRIOR FILING DATE: 2000-05-02
; NUMBER OF SEQ ID NOS: 182
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-783-950-25

Query Match 61.5%; Score 40; DB 16; Length 13;
Best Local Similarity 66.7%; Pred. No. 11;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 SWDSSGTHP 10
:|||||
Db 2 AWDSSSDHP 10

RESULT 15

US-10-424-599-243244
; Sequence 243244, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 243244
; LENGTH: 90
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_61678C.1.pep
US-10-424-599-243244

Query Match 61.5%; Score 40; DB 15; Length 90;
Best Local Similarity 66.7%; Pred. No. 74;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 NSWDSSGTH 9
:|||||
Db 11 NSWDEGGDH 19

Search completed: April 28, 2005, 19:02:14
Job time : 69.2344 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2005, 18:05:55 ; Search time 15.125 Seconds
(without alignments)
69.976 Million cell updates/sec

Title: US-10-088-639A-2_COPY_88_98
Perfect score: 65
Sequence: 1 NSWDSSGTHPV 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	64.6	130	2	Y0P6 chaperone syc
2	41	63.1	98	2	probable HD super
3	41	63.1	339	2	hypothetical prote
4	41	63.1	456	2	hypothetical prote
5	40	61.5	18	2	Tcr C gamma 1 chai
6	40	61.5	110	2	T-cell receptor ga
7	40	61.5	120	2	T-cell receptor ga
8	40	61.5	132	2	Ig lambda chain V
9	40	61.5	629	2	hypothetical prote
10	40	61.5	914	2	probable myrosinas
11	40	61.5	988	2	myrosinase-binding
12	39	60.0	201	2	protein F26H11.1
13	39	60.0	270	2	lectin - garden pe
14	39	60.0	402	2	hypothetical prote
15	39	60.0	652	2	probable CHP-rich
16	39	60.0	995	2	hypothetical prote
17	39	60.0	1336	2	hypothetical prote
18	38	58.5	295	2	hypothetical prote
19	38	58.5	451	1	NifH oxidoreductas
20	38	58.5	476	2	endoglycosylcerami
21	38	58.5	567	2	oligopeptide ABC t
22	38	58.5	1099	2	myosin VIII, ZMM3
23	38	58.5	1290	2	leucocyte common a
24	38	58.5	1369	2	hypothetical prote
25	38	58.5	1452	1	protein-tyrosine-p
26	38	58.5	1452	1	protein-tyrosine-p
27	38	58.5	1797	2	cell surface glyco
28	38	58.5	1797	2	leukocyte antigen-
29	38	58.5	1898	2	leukocyte antigen-

30 37 56.9 108 2 S47184
31 37 56.9 109 2 S19663
32 37 56.9 125 2 C83704
33 37 56.9 271 2 B86918
34 37 56.9 294 2 JCS000
35 37 56.9 295 2 S52791
36 37 56.9 295 2 G02924
37 37 56.9 295 2 A55451
38 37 56.9 295 2 JCS2523
39 37 56.9 295 2 JCS249
40 37 56.9 295 2 S52794
41 37 56.9 295 2 S52794
42 37 56.9 295 2 JCS248
43 37 56.9 295 2 G01843
44 37 56.9 330 2 JCS935
45 37 56.9 372 2 A53050

ALIGNMENTS

RESULT 1

T43606
Y0P6 chaperone sycE - Yersinia pestis plasmid pCD1
C:Species: Yersinia pestis
C>Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004
C:Accession: T43606; T42911
R:Hu P.; Elliott, J.; McCready, P.; Skowronski, E.; Garnes, J.; Kobayashi, A.; Brubaker, J.
J. Bacteriol. 180, 5192-5202, 1998
A:Title: Structural organization of virulence-associated plasmids of Yersinia pestis.
A:Reference number: 222578; MUID:98422474; PMID:9748454
A:Accession: T43606
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-130 <HUP>
A:Cross-references: UNIPROT:P31491; EMBL:AF053946; NID:g2996222; PIDN:AAC62588.1; PFI:92
A:Experimental source: strain KIM
R:Perry, R.D.; Straley, S.C.; Fetherston, J.D.; Rose, D.J.; Gregor, J.; Blattner, F.R.
Infect. Immun. 66, 4611-4623, 1998
A:Title: DNA sequencing and analysis of the low-Ca2+-response plasmid pCD1 of Yersinia f
A:Reference number: 222273; MUID:98427122; PMID:9746557
A:Accession: T42911
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-130 <PER>
A:Cross-references: EMBL:AF074612; NID:g3822037; PIDN:AAC69819.1; PID:g3822099
A:Experimental source: strain KIMS
C:Genetics:
A:Gene: sycE
A:Genome: plasmid pCD1
C:Superfamily: Yersinia pestis plasmid pCD1 hypothetical protein sycE

Query Match 64.6%; Score 42; DB 2; Length 130;
Best Local Similarity 70.0%; Pred. No. 4.4;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 SWDSSGTHPV 11
||| |
DB 79 SWDEVGGHPV 88

RESULT 2

F97245
probable HD superfamily hydrolase [imported] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C>Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
C:Accession: F97245
R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markatova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cl
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: F97245

A:Introns: 15/1
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: T-cell receptor

Query Match 61.5%; Score 40; DB 2; Length 120;
Best Local Similarity 77.8%; Pred. No. 8.9;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 WDSSGTHPV 11
|||||
Db 112 WDSSGFHKV 120

RESULT 8

S09713
Ig lambda chain V region - human
C:Species: Homo sapiens (man)
C:Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C:Accession: S09713
R: Hughes-Jones, N.C.; Bye, J.M.; Beale, D.; Coadwell, J.
Biochem. J. 268, 135-140, 1990
A:Title: Nucleotide sequences and three-dimensional modelling of the VH and VL domains of
A:Reference number: S09710; MUID:90262535; PMID:2111699
A:Accession: S09713
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-132 <HUG>
C:Cross-references: UNIPROT:O8N355
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotrimer; immunoglobulin
F:34-110/Domain: immunoglobulin homology <IRMM>

Query Match 61.5%; Score 40; DB 2; Length 132;
Best Local Similarity 75.0%; Pred. No. 9.8;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 WDSSGTHP 10
|||||
Db 111 WDSSSAHP 118

RESULT 9

T38214
hypothetical protein SPAC22H10.11c - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T38214
R:Grishchuk, K.; McIntosh, J.R.; Devlin, K.; Churcher, C.; Barrell, B.G.; Rajandream, M.
submitted to the EMBL Data Library, February 1996
A:Reference number: Z21778
A:Accession: T38214
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-629 <DEV>
C:Cross-references: UNIPROT:Q10304; EMBL:Z69730; PIDN:CAA93611.1; GSPDB:GN000066; SPDB:SE
A:Experimental source: strain 972h-; cosmid c22H10
C:Genetics:
A:Gene: SPDB:SPAC22H10.11c
A:Map position: 1
C:Superfamily: Schizosaccharomyces pombe hypothetical protein SPAC22H10.11c

Query Match 61.5%; Score 40; DB 2; Length 629;
Best Local Similarity 60.0%; Pred. No. 52;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 NSWDSSGTHP 10
|||||
Db 100 NSWKSSSHP 109

RESULT 10

T08081
probable myrosinase-binding protein - rape

N:Alternate names: jasmonate inducible protein
C:Species: Brassica napus (rape)
C:Date: 21-May-1999 #sequence_revision 21-May-1999 #text_change 09-Jul-2004
C:Accession: T08081
R:Geshi, N.; Brandt, A.
Planta 204, 295-304, 1998
A:Title: Two jasmonate-inducible myrosinase-binding proteins from Brassica napus L. seed
A:Reference number: Z16340; MUID:98192006; PMID:9530873
A:Accession: T08081
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-914 <GES>
C:Cross-references: UNIPROT:P93659; EMBL:X11482; NID:e1023100; PIDN:CAA72270.1; PID:e30
A:Experimental source: cv. Global; isolate a4; seedlings
A:Note: jasmonate inducible

Query Match 61.5%; Score 40; DB 2; Length 914;
Best Local Similarity 66.7%; Pred. No. 78;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 NSWDSSGTH 9
|||||
Db 683 NQWDDSGDH 691

RESULT 11

T08102
myrosinase-binding protein (clone indmbp2) - rape
C:Species: Brassica napus (rape)
C:Date: 21-May-1999 #sequence_revision 21-May-1999 #text_change 09-Jul-2004
C:Accession: T08102
R:Taipalensuu, J.; Eriksson, S.J.; Rask, L.
Eur. J. Biochem. 250, 680-688, 1997
A:Title: The myrosinase binding protein from Brassica napus seeds possesses lectin acti
A:Reference number: Z16355; MUID:98121188; PMID:9461290
A:Accession: T08102
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-988 <TAI>
C:Cross-references: UNIPROT:P93065; EMBL:Y09437; NID:g1711295; PIDN:CAA70587.1; PID:g17
A:Experimental source: cv. Hanna; leaves

Query Match 61.5%; Score 40; DB 2; Length 988;
Best Local Similarity 66.7%; Pred. No. 85;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 NSWDSSGTH 9
|||||
Db 699 NQWDDSGDH 707

RESULT 12

A88363
protein F26H11.1 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: A88363
R:Anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolo
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_el
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; an
A:Accession: A88363
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-201 <STO>
C:Cross-references: GB:chr_II; PIDN:CAB04194.1; PID:g3876448; GSPDB:GN000020; CBSP:F26H11
A:Note: predicted using Genefinder
C:Genetics:
A:Gene: F26H11.1
A:Map position: 2

Query Match 60.0%; Score 39; DB 2; Length 201;
Best Local Similarity 66.7%; Pred. No. 23;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 2 NSWSSGTHP 10
DB 101 SWSNSGFHP 109
RESULT 13
T06528
lectin - garden pea
C;Species: Pisum sativum (garden pea)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: T06528
R;Kardailsky, I.V.; Sherrier, D.J.; Brewin, N.J.
Plant Physiol. 111, 49-60, 1995
A;Title: Identification of a new pea gene PsNlec1 that encodes a lectin-like glycoprotein
A;Reference number: 215736
A;Accession: T06528
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-270 <K>
A;Cross-references: UNIPROT:Q40987; EMBL:U31981; NID:G974303; PIDN:AAC49367.1; PID:G9743
A;Experimental source: cv. Wisconsin Perfection
C;Genetics:
A;Gene: Nlec1
C;Superfamily: plant lectin
Query Match 60.0%; Score 39; DB 2; Length 270;
Best Local Similarity 66.7%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 NSWSSGTH 9
DB 162 NSWDPEGRH 170
RESULT 14
S74525
hypothetical protein slr0731 - Synechocystis sp. (strain PCC 6803)
C;Species: Synechocystis sp.
A;Variety: PCC 6803
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C;Accession: S74525
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
Okada, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
s.
A;Reference number: S74322; MUID:97061201; PMID:8905231
A;Accession: S74525
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-402 <KAN>
A;Cross-references: UNIPROT:P72675; EMBL:D90899; GB:AB001339; NID:g1651650; PIDN:BAAL667
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C;Superfamily: Synechocystis hypothetical protein slr0731
Query Match 60.0%; Score 39; DB 2; Length 402;
Best Local Similarity 66.7%; Pred. No. 48;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 NSWSSGTH 9
DB 197 DSWSSSTH 205
RESULT 15
G85024
probable CHP-rich zinc finger protein [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004

C;Accession: G85024
R;anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
Nature 402, 769-777, 1999
A;Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A;Reference number: AB5001; MUID:20083488; PMID:10617198
A;Accession: G85024
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-652 <STO>
A;Cross-references: UNIPROT:Q9SVJ0; GB:NC_001269; NID:G7268577; PIDN:CAB80686.1; GSPDB:G
C;Genetics:
A;Gene: AT4g01930
A;Map position: 4
Query Match 60.0%; Score 39; DB 2; Length 652;
Best Local Similarity 54.5%; Pred. No. 81;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 1 NSWSSGTHPV 11
DB 393 NKWCSACTHPI 403
Search completed: April 28, 2005, 18:25:53
Job time : 17.125 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 27, 2004, 12:07:29 ; Search time 23 Seconds
(without alignments)
24.903 Million cell updates/sec

Title: US-10-088-639A-2_COPY_88_98
Perfect score: 65
Sequence: 1 NSWDSSGTHPV 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	42	64.6	130	1 YERA_YEREN	P31490 yersinia en
2	42	64.6	130	1 YERA_YERPE	P31491 yersinia pe
3	40	61.5	629	1 YD4B_SCHPO	Q10304 schizosacch
4	38	58.5	451	1 YJ66_YEAST	P47051 saccharomyc
5	38	58.5	476	1 NIFE_RHIME	Q92122 rhizobium m
6	38	58.5	496	1 NIFE_RHISN	P56673 rhizobium s
7	38	58.5	511	1 MATK_BROIN	Q9muzz bromus iner
8	38	58.5	1452	1 PTFM_HUMAN	P28827 homo sapien
9	38	58.5	1452	1 PTFM_MOUSE	P28828 mus musculu
10	37	56.9	88	1 DHH_BRARE	P79729 brachydanio
11	37	56.9	294	1 SUPP_BOVIN	P50227 bos taurus
12	37	56.9	295	1 SUP1_HUMAN	P50225 homo sapien
13	37	56.9	295	1 SUP2_HUMAN	P50226 homo sapien
14	37	56.9	295	1 SUPM_HUMAN	P50224 h monoamine
15	37	56.9	295	1 SUPP_MACFA	P52846 macaca fasc
16	37	56.9	330	1 EXL2_HUMAN	Q9ubq6 homo sapien
17	37	56.9	330	1 EXL2_MOUSE	Q9se89 mus musculu
18	37	56.9	375	1 PELB_ERWCH	P04959 erwinia chr
19	37	56.9	375	1 PELC_ERWCH	P11073 erwinia chr
20	37	56.9	666	1 ENV_MLVHO	P21436 homulv muri
21	37	56.9	1302	1 GPRS_DROME	Q61366 drosophila
22	36	55.4	234	1 YV04_MYCTU	Q50721 mycobacteri
23	36	55.4	250	1 AGL9_ARADE	Q38694 aranda debo
24	36	55.4	406	1 P152_HUMAN	P48426 homo sapien
25	36	55.4	406	1 P153_HUMAN	P53807 homo sapien
26	36	55.4	595	1 SNX9_HUMAN	Q9Y5x1 homo sapien
27	36	55.4	665	1 ENV_MLVMO	P03385 moloney mur
28	36	55.4	671	1 HPS4_MOUSE	Q99k97 mus musculu
29	36	55.4	755	1 BPHY_DEIRA	Q9rza4 deinococcus
30	36	55.4	811	1 RPOP_NEUIN	P33541 neurospora
31	36	55.4	831	1 RPOP_GELSP	O03685 gelatinospo
32	36	55.4	1036	1 Y946_ARCFU	O29316 archaeoglob
33	35	53.8	224	1 HAD2_PSEPU	Q59728 pseudomonas

P24070	pseudomonas
P55890	salmomella
Q34943	lumbriacus t
Q99ur8	staphylococ
Q8cpj7	staphylococ
Q9hkm2	thermoplasm
P77883	lactobacill
Q97b29	thermoplasm
Q60252	lactobacill
Q9vnb5	drosophila
Q7vr80	candidate
Q8yp12	anabaena sp
Q9xkf6	helicobacte
P56008	helicobacte
Q96a23	homo sapien
Q91vh2	mus musculu
P71353	haemophilus
Q8K320	mus musculu
Q63258	rattus norv
Q61738	mus musculu
P01714	homo sapien
P41309	didelphis m
Q13291	homo sapien
P80748	streptococ
P80598	streptococ
P29609	streptococ
P20250	pseudechis
P20254	pseudechis
P20251	pseudechis
P55273	homo sapien
Q60773	mus musculu
Q63148	rattus norv
P52198	homo sapien
Q99ym5	mus musculu
P81371	vatairea ma
Q58624	methanococ
P27859	escherichia
P21444	feline leuk
Q92826	homo sapien
Q12437	aspergillus
Q39527	cladarratis
Q8eux9	mycoplasma
Q8g89	fusobacteri
O28293	methanobact
Q8qph9	streptococ
Q914t8	enterococcu
Q9cf79	lactococcu
Q814n6	lactococcu
Q8p191	streptococ
Q9A0C8	streptococ
O04135	malus domes
O04134	malus domes
Q60078	schizosacch
Q9p37	xyella fas
Q839V7	enterococcu
P24423	azotobacter
Q91bg7	homo sapien
Q6187	rattus norv
P43662	salmomella
P23059	bovine leuk
Q03361	bovine leuk
O60309	homo sapien
Q920e2	mus musculu
Q9b2x0	homo sapien
P24004	saccharomyc
P15436	saccharomyc
Q9cpk3	pasteurella
P21543	paenibacill
Q9upq9	homo sapien

RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague."
RL Nature 413:523-527(2001).
CC -!- FUNCTION: POSITIVE REGULATOR OF YOPE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M34279; AAA27671.1; -.
DR EMBL; AF074612; AAC69819.1; -.
DR EMBL; AF053946; AAC62588.1; -.
DR EMBL; AL117189; CAB54882.1; -.
DR PIR; T43606; T43606.
DR PDB; 1UYA; 13-MAR-02.
DR InterPro; IPR005416; SYCE_chap.
DR PRINTS; PR01596; SYCECHAPRONE.
KW Virulence; Plasmid; Transcription regulation; Activator;
KW Complete proteome; 3D-structure.
SQ SEQUENCE 130 AA; 14650 MW; 41A12BB29831CA5A CRC64;

Query Match 64.6%; Score 42; DB 1; Length 130;
Best Local Similarity 70.0%; Pred. No. 1.9;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 SWDSGTHPV 11
DB 79 SWDEGVGHPV 88

RESULT 3
YD4B_SCHPO
ID YD4B SCHPO STANDARD; PRT; 629 AA.
AC Q10304;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein C2H10.11c in chromosome I.
GN SPAC22H10.11c.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Williams R., Rejandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leathers S., McDonald S., Mclean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volktaert G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanterre E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gallard C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,

RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RL "The genome sequence of Schizosaccharomyces pombe."
RT Nature 415:871-880(2002).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z69730; CAA93611.1; -.
DR PIR; T38214; T38214.
DR GeneDB_Spombe; SPAC22H10.11c; -.
KW Hypothetical protein.
FT DOMAIN 201 208 POLY-GLU.
SQ SEQUENCE 629 AA; 71266 MW; A32FA33FEE27605D CRC64;

Query Match 61.5%; Score 40; DB 1; Length 629;
Best Local Similarity 60.0%; Pred. No. 24;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 NSWDSSGTHP 10
DB 100 NSWKSSSHP 109

RESULT 4
YJ66 YEAST
ID YJ66 YEAST STANDARD; PRT; 451 AA.
AC P47051.
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Hypothetical 52.1 kDa protein in MTR4-GYP6 intergenic region.
GN YJL046W OR J1171.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Pohl T.M., Aljinovic G.;
RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: SOME, TO BACTERIAL LIPOATE-PROTEIN LIGASE A (LPLA).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z49321; CAA89337.1; -.
DR PIR; S56818; S56818.
DR GeneOnline; 141660; -.
DR SGD; S0003582; YJL046W.
DR InterPro; IPR004143; BPL_LipA_LipB.
DR InterPro; IPR004562; Lipoyltrans.
DR Pfam; PF03099; BPL_LipA_LipB; 1.
DR TIGRFAMs; TIGR00545; lipoyltrans; 1.
KW Hypothetical protein.
SQ SEQUENCE 451 AA; 52085 MW; 637673A018B865AE CRC64;

Query Match 58.5%; Score 38; DB 1; Length 451;
Best Local Similarity 66.7%; Pred. No. 37;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 WDSGTHPV 11

```

Db      333 WESSGVHSV 341
|:|||||
RESULT 5
NIFE_RHIME
ID NIFE_RHIME STANDARD; PRT; 476 AA.
AC Q92L2;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Nitrogenase iron-molybdenum cofactor biosynthesis protein nife.
GN NIFE OR RA0452 OR SNA0830.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OG Plasmid pSymA (megaplasmid 1).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21396509; PubMed=11481432;
RA Barnett M.J., Fisher R.P., Jones T., Komp C., Abola A.P.,
RA Barloy-Hubler F., Bowser L., Capela D., Galibert F., Gouzy J.,
RA Gurjal M., Hong A., Huizar L., Hyman R.W., Kahn D., Kahn M.L.,
RA Kalman S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,
RA Yeh K.-C., Davis R.W., Federspiel N.A., Long S.R.;
RT "Nucleotide sequence and predicted functions of the entire
RT Sinorhizobium meliloti pSymA megaplasmid.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888(2001).
CC -!- FUNCTION: This protein may play a role in the biosynthesis of the
CC prosthetic group of nitrogenase (FeMo cofactor).
CC -!- PATHWAY: Fe-Mo cofactor biosynthesis.
CC -!- SIMILARITY: BELONGS TO THE NIFD/NIFK/NIFE/NIFN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB007236; AAK65110.1; -.
CC PIR; D95318; D95318.
CC InterPro; IPR005973; Nife.
CC InterPro; IPR000318; Nitrogense compl.
CC InterPro; IPR000510; Oxred nitrogensel.
CC Pfam; PF00148; Oxidored nitro; 1.
CC TIGRFAMs; TIGR01283; nife; 1.
CC PROSITE; PS00699; NITROGENASE_1_1; 1.
CC PROSITE; PS00090; NITROGENASE_1_2; 1.
CC Nitrogen fixation; Plasmid.
CC SEQUENCE 476 AA; 52358 MW; 83CE94B4B14C713 CRC64;
Query Match 58.5%; Score 38; DB 1; Length 476;
Best Local Similarity 75.0%; Pred. No. 39;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 NSWDSSGT 8
|:|||||
DB 73 NSWDNRGT 80
RESULT 6
NIFE_RHISN
ID NIFE_RHISN STANDARD; PRT; 496 AA.
AC P5673;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Nitrogenase iron-molybdenum cofactor biosynthesis protein nife.
GN NIFE OR Y4VN.
OS Rhizobium sp. (strain NGR234).
OG Plasmid sym pNGR234a.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.
OX NCBI_TaxID=394;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97305956; PubMed=9163424;
RA Freiberg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A.,
RA Perret X.;
RT "Molecular basis of symbiosis between Rhizobium and legumes.";
RL Nature 387:394-401(1997).
CC -!- FUNCTION: This protein may play a role in the biosynthesis of the
CC prosthetic group of nitrogenase (FeMo cofactor).
CC -!- PATHWAY: Fe-Mo cofactor biosynthesis.
CC -!- SIMILARITY: BELONGS TO THE NIFD/NIFK/NIFE/NIFN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB000102; AAB91902.1; -.
CC HSP; P07328; 3MIN.
CC InterPro; IPR005973; Nife.
CC InterPro; IPR000318; Nitrogense compl.
CC InterPro; IPR000510; Oxred nitrogensel.
CC Pfam; PF00148; Oxidored nitro; 1.
CC TIGRFAMs; TIGR01283; nife; 1.
CC PROSITE; PS00699; NITROGENASE_1_1; 1.
CC PROSITE; PS00090; NITROGENASE_1_2; 1.
CC Nitrogen fixation; Plasmid.
CC SEQUENCE 496 AA; 54793 MW; D78472D8F5410A3A CRC64;
Query Match 58.5%; Score 39; DB 1; Length 496;
Best Local Similarity 75.0%; Pred. No. 41;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 NSWDSSGT 8
|:|||||
DB 73 NSWDNRGT 80
RESULT 7
MATK_BROIN
ID MATK_BROIN STANDARD; PRT; 511 AA.
AC Q9WU23;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Maturase K (intron maturase).
GN MATK.
OS Bromus inermis (Smooth brome grass).
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC Bromaceae; Bromus.
OX NCBI_TaxID=15371;
RN [1]
RP SEQUENCE FROM N.A.
RA Hilu K.W., Alice L.A., Liang H.;
RT "Phylogeny of Poaceae inferred from matK sequences.";
RL Ann. Mo. Bot. Gard. 86:835-851(1999).
CC -!- FUNCTION: Probably assists in splicing chloroplast group II
CC introns (By similarity).
CC -!- SIMILARITY: Belongs to the intron maturase family 2. Matk
CC subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AF164398; AAF66185.1; -;
 DR InterPro; IPR000442; Intron mature2.
 DR InterPro; IPR002866; MatK.N.
 DR Pfam; PF01348; Intron maturas2; 1.
 DR Pfam; PF01824; MatK.N; 1.
 KW mRNA processing; Chloroplast.
 SQ SEQUENCE 511 AA; 61572 MW; 088497F25BE52C88 CRC64;

Query Match 58.5%; Score 38; DB 1; Length 511;
 Best Local Similarity 60.0%; Pred. No. 42;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 1 NSWSSGTHP 10
 DB 69 NFWDSNVNHP 78

RESULT 8

PTPM HUMAN
 ID PTPM HUMAN STANDARD; PRT; 1452 AA.
 AC P28827;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Receptor-type protein-tyrosine phosphatase mu precursor (EC 3.1.3.48)
 DE (R-PTP-mu).
 GN PTPRM OR PTPRL1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92008644; PubMed=1655529;
 RA Gebbink M.P.B.G., van Eeten I., Hateboer G., Suijkerbuijk R.,
 RA Beijersbergen R., Geurts van Kessel A., Moolenaar W.H.;
 RT "Cloning, expression and chromosomal localization of a new putative
 RT receptor-like protein tyrosine phosphatase.";
 RL FEBS Lett. 290:123-130(1991).
 RN [2]
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 879-1156.
 RX MEDLINE=98010572; PubMed=9346878;
 RA Hoffmann K.M., Tonks N.K., Barford D.;
 RT "The crystal structure of domain 1 of receptor protein-tyrosine
 RL J. Biol. Chem. 272:27505-27508(1997).
 CC -!- FUNCTION: May play a key role in signal transduction and growth
 CC control.
 CC -!- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
 CC tyrosine + phosphate.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- SIMILARITY: Contains 1 MAM domain.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC -!- SIMILARITY: Contains 4 fibronectin type III domains.
 CC -!- SIMILARITY: Contains 2 protein-tyrosine phosphatase domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; X58288; CAA41226.1; -;
 DR PIR; S17669; S17669.
 DR PDB; 1RPM; 01-APR-98.

DR Genew; HGNC:9675; PTPRM.
 DR MIM; 176888; -;
 DR GO; GO:0005887; C:integral to plasma membrane; TAS.
 DR GO; GO:0005001; P:transmembrane receptor protein tyrosine pho. . .; TAS.
 DR GO; GO:0006470; P:protein amino acid dephosphorylation; TAS.
 DR InterPro; IPR008985; ConA-like lec_gl.
 DR InterPro; IPR008957; FN III-like.
 DR InterPro; IPR003961; FN III.
 DR InterPro; IPR007110; IG-like.
 DR InterPro; IPR003599; IG.
 DR InterPro; IPR000998; MAM domain.
 DR InterPro; IPR000387; TYR_phosphatase.
 DR InterPro; IPR000242; Tyr_PP.
 DR Pfam; PF00041; fn3; 3.
 DR Pfam; PF00047; ig; 1.
 DR Pfam; PF00629; MAM; 1.
 DR Pfam; PF00102; Y_phosphatase; 2.
 DR PRINTS; PR00020; MAMDOMAIN.
 DR PRINTS; PR00700; PRTYPHPTASE.
 DR SMART; SM00060; FN3; 2.
 DR SMART; SM00409; IG; 1.
 DR SMART; SM00137; MAM; 1.
 DR SMART; SM00194; PTPC; 2.
 DR PROSITE; PS50835; IG_LIKE; 1.
 DR PROSITE; PS00740; MAM_1; 1.
 DR PROSITE; PS50060; MAM_2; 1.
 DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 2.
 DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 2.
 DR PROSITE; PS50055; TYR_PHOSPHATASE_PTP; 2.
 KW Hydrolase; Receptor; Glycoprotein; Signal; Transmembrane;
 KW Immunoglobulin domain; Repeat; 3D-structure.
 FT SIGNAL 1 20 POTENTIAL.
 FT CHAIN 21 1452 RECEPTOR-TYPE PROTEIN-TYROSINE
 FT PTPRM PHOSPHATASE MU.
 FT DOMAIN 21 742 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 743 764 POTENTIAL.
 FT DOMAIN 765 1452 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 22 184 MAM.
 FT DOMAIN 186 277 IG-LIKE C2-TYPE.
 FT DOMAIN 287 374 FIBONECTIN TYPE-III 1.
 FT DOMAIN 383 466 FIBONECTIN TYPE-III 2.
 FT DOMAIN 486 571 FIBONECTIN TYPE-III 3.
 FT DOMAIN 589 671 FIBONECTIN TYPE-III 4.
 FT DOMAIN 923 1153 PROTEIN-TYROSINE PHOSPHATASE 1.
 FT DOMAIN 1213 1447 PROTEIN-TYROSINE PHOSPHATASE 2.
 FT ACT_SITE 1095 1095 PHOSPHOCYSTEINE INTERMEDIATE (BY
 FT ACT_SITE 1389 1389 SIMILARITY).
 FT ACT_SITE 1389 1389 PHOSPHOCYSTEINE INTERMEDIATE (BY
 FT DISULFID 206 260 SIMILARITY).
 FT CARBOHYD 72 72 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 92 92 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 131 131 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 249 249 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 406 406 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 414 414 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 454 454 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 534 534 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 544 544 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 598 598 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 651 651 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 681 681 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT STRAND 880 881
 FT TURN 882 884
 FT HELIX 885 905
 FT TURN 906 906
 FT TURN 916 919
 FT HELIX 921 926
 FT TURN 930 931
 FT HELIX 936 938
 FT STRAND 939 941
 FT TURN 946 947
 FT TURN 949 952

FT STRAND 953 961
 FT TURN 962 963
 FT STRAND 964 971
 FT HELIX 976 978
 FT TURN 979 988
 FT TURN 989 990
 FT STRAND 993 996
 FT STRAND 1001 1002
 FT TURN 1003 1004
 FT STRAND 1005 1006
 FT STRAND 1016 1019
 FT TURN 1020 1021
 FT STRAND 1022 1032
 FT TURN 1033 1034
 FT STRAND 1035 1044
 FT TURN 1045 1046
 FT STRAND 1051 1058
 FT HELIX 1071 1083
 FT TURN 1086 1087
 FT STRAND 1091 1094
 FT HELIX 1100 1118
 FT STRAND 1120 1121
 FT HELIX 1123 1133
 FT TURN 1135 1136
 FT HELIX 1141 1155
 SQ SEQUENCE 1452 AA; 163633 MW; 08175D3595A6C7E0 CRC64;

 Query Match 58.5%; Score 38; DB 1; Length 1452;
 Best Local Similarity 55.6%; Pred. No. 1.3e+02;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

 QY 2 SWDSSGTHP 10
 Db 434 SWDTENSHP 442
 |||: |||

 RESULT 9
 PTPM_MOUSE STANDARD; PRT; 1452 AA.
 AC P28828;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Receptor-type protein-tyrosine phosphatase mu precursor (EC 3.1.3.48)
 DE (R-PTP-mu).
 GN PTPRM.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RX MEDLINE=92008644; PubMed=1655529;
 RA Gebink M.F.B.G., van Eeten I., Hateboer G., Suijkerbuijk R.,
 RA Beijersbergen R., Geurts van Kessel A., Moolenaar W.H.;
 RT "Cloning, expression and chromosomal localization of a new putative
 RT receptor-like protein tyrosine phosphatase.";
 RL FEBS Lett. 290:123-130(1991).
 CC -!- FUNCTION: May play a key role in signal transduction and growth
 CC control.
 CC -!- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
 CC tyrosine + phosphate.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- TISSUE SPECIFICITY: Most abundant in lung, less in brain and
 CC heart.
 CC -!- SIMILARITY: Contains 1 MAM domain.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC -!- SIMILARITY: Contains 4 fibronectin type III domains.
 CC -!- SIMILARITY: Contains 2 protein-tyrosine phosphatase domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; X58287; CAA41225.1; -;
 DR PIR; S17670; S17670.
 DR HSSP; P28827; IRPM.
 DR MGD; MG1:102694; Ptpm.
 DR InterPro; IPR008985; ConA like lec_gl.
 DR InterPro; IPR008957; FN_III-like.
 DR InterPro; IPR003961; FN_III.
 DR InterPro; IPR007110; IG-like.
 DR InterPro; IPR003599; IG.
 DR InterPro; IPR000998; MAM_domain.
 DR InterPro; IPR000387; TYR_phosphatase.
 DR InterPro; IPR000242; Tyr_PP.
 DR Pfam; PF00041; fn3; 3.
 DR Pfam; PF00047; ig; 1.
 DR Pfam; PF00629; MAM; 1.
 DR Pfam; PF0102; Y_phosphatase; 2.
 DR PRINTS; PR00020; MAMDOMAIN.
 DR PRINTS; PR00700; PRTYPHPTASE.
 DR SMART; SM00060; FN3; 2.
 DR SMART; SM00409; IG; 1.
 DR SMART; SM00137; MAM; 1.
 DR SMART; SM00194; PTPC; 2.
 DR PROSITE; PS0835; IG_LIKE; 1.
 DR PROSITE; PS00740; MAM_1; 1.
 DR PROSITE; PS0060; MAM_2; 1.
 DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 2.
 DR PROSITE; PS00056; TYR_PHOSPHATASE_2; 2.
 DR PROSITE; PS0055; TYR_PHOSPHATASE_PTP; 2.
 KW Hydrolase; Receptor; Glycoprotein; Signal; Transmembrane;
 KW Immunoglobulin domain; Repeat.
 FT SIGNAL 1 20
 FT CHAIN 21 1452
 FT
 FT DOMAIN 21 742
 FT TRANSMEM 743 764
 FT DOMAIN 765 1452
 FT DOMAIN 22 184
 FT DOMAIN 186 277
 FT DOMAIN 287 374
 FT DOMAIN 383 466
 FT DOMAIN 486 571
 FT DOMAIN 589 671
 FT DOMAIN 923 1153
 FT DOMAIN 1213 1447
 FT ACT_SITE 1095 1095
 FT ACT_SITE 1389 1389
 FT
 FT DISULFID 206 260
 FT CARBOHYD 72 72
 FT CARBOHYD 92 92
 FT CARBOHYD 131 131
 FT CARBOHYD 249 249
 FT CARBOHYD 406 406
 FT CARBOHYD 414 414
 FT CARBOHYD 454 454
 FT CARBOHYD 534 534
 FT CARBOHYD 544 544
 FT CARBOHYD 598 598
 FT CARBOHYD 651 651
 FT CARBOHYD 681 681
 SQ SEQUENCE 1452 AA; 163594 MW; F99D0FC67922CF1E CRC64;

 Query Match 58.5%; Score 38; DB 1; Length 1452;
 Best Local Similarity 55.6%; Pred. No. 1.3e+02;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

QY 2 SWDSSGTHP 10
Db 434 SWDTSNHP 442

RESULT 10
DHH_BRARE STANDARD; PRT; 88 AA.
AC P79729; Q9YGUJ3;
DT 15-JUL-1999 (Rel. 38, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Desert hedgehog protein (DHH) (Fragment).
GN DHH.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99051425; PubMed=9831563;
RA Amores A., Force A., Yan Y.-L., Wang Y.-L., Fritz A., Prince V.,
RA Ho R., Amemiya C., Langeland J., Westerfield M., Ekker M.,
RA Postlethwait J.,
RT "Zebrafish hox clusters and vertebrate genome evolution.";
RL Science 282:1711-1714(1998).
RN [2]
RP SEQUENCE OF 14-71 FROM N.A.
RC TISSUE=Muscle;
RX MEDLINE=97075114; PubMed=8917540;
RA Zardoya R., Abouheif E., Meyer A.;
RT "Evolutionary analyses of hedgehog and Hoxd-10 genes in fish species
closely related to the zebrafish.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:13036-13041(1996).
RN [1]
RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RX TISSUE=Tracheobronchial;
RX MEDLINE=96003918; PubMed=7575456;
RA Schaus S.J., Henry T., Palmatier R., Halvorson L., Dannenbring R.,
RA Beckmann J.D.;
RT "Characterization of bovine tracheobronchial phenol sulphotransferase
cDNA and detection of mRNA regulation by cortisol.";
RL Biochem. J. 311:209-217(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA Henry T., Kliewer B., Palmatier R., Ulphani J., Beckmann J.D.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 24-57 FROM N.A.
RA Nonneman D.J., Shibuya H., Johnson G.S.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Catalyzes the sulfate conjugation of catecholamines and
of phenolic drugs.
CC -!- CATALYTIC ACTIVITY: 3'-phosphoadenylylsulfate + a phenol =
adenosine 3',5'-bisphosphate + an aryl sulfate.
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- TISSUE SPECIFICITY: Distal lung parenchyma.
CC -!- SIMILARITY: Belongs to the sulfotransferase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U35253; AAA85510.1; -
CC EMBL; U34753; AAC48677.1; -
CC EMBL; L33828; AAA56789.1; -
CC PIR; JC5000; JC5000.
CC HSSP; P50224; 1CUM.
CC InterPro; IPR000863; Sulfotransferase.
CC Pfam; PF00685; Sulfotransfer; 1.

Query Match 56.9%; Score 37; DB 1; Length 88;
Best Local Similarity 55.6%; Pred. No. 9.3; Indels 0; Gaps 0;
Matches 5; Conservative 1; Mismatches 3;

QY 2 SWDSSGTHP 10
Db 28 AWDEGHP 36

RESULT 11
SUPP BOVIN STANDARD; PRT; 294 AA.
AC P50227;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Phenol-sulfating phenol sulfotransferase (EC 2.8.2.1) (P-PST).
GN STP.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RX TISSUE=Tracheobronchial;
RX MEDLINE=96003918; PubMed=7575456;
RA Schaus S.J., Henry T., Palmatier R., Halvorson L., Dannenbring R.,
RA Beckmann J.D.;
RT "Characterization of bovine tracheobronchial phenol sulphotransferase
cDNA and detection of mRNA regulation by cortisol.";
RL Biochem. J. 311:209-217(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA Henry T., Kliewer B., Palmatier R., Ulphani J., Beckmann J.D.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 24-57 FROM N.A.
RA Nonneman D.J., Shibuya H., Johnson G.S.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Catalyzes the sulfate conjugation of catecholamines and
of phenolic drugs.
CC -!- CATALYTIC ACTIVITY: 3'-phosphoadenylylsulfate + a phenol =
adenosine 3',5'-bisphosphate + an aryl sulfate.
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- TISSUE SPECIFICITY: Distal lung parenchyma.
CC -!- SIMILARITY: Belongs to the sulfotransferase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U35253; AAA85510.1; -
CC EMBL; U34753; AAC48677.1; -
CC EMBL; L33828; AAA56789.1; -
CC PIR; JC5000; JC5000.
CC HSSP; P50224; 1CUM.
CC InterPro; IPR000863; Sulfotransferase.
CC Pfam; PF00685; Sulfotransfer; 1.

```

DR ProDom: PD001218; Sulfotransferase; 1.
 KW Catecholamine metabolism; Steroid metabolism; Transferase.
 FT BINDING 259 265 FAPS BINDING SITE (POTENTIAL).
 SQ SEQUENCE 294 AA; 34017 MW; 8ADEE67D47E69737 CRC64;

Query Match 56.9%; Score 37; DB 1; Length 294;
 Best Local Similarity 66.7%; Pred. No. 34;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 3 WDSGTHFV 11
 | : | | | |
 Db 180 WELSHTHFV 188

RESULT 12
 SUP1 HUMAN
 ID SUP1 HUMAN STANDARD; PRT; 295 AA.
 AC P50225; Q92818; Q9BVU6; Q9UGG7;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Phenol-sulfating phenol sulfotransferase 1 (P-PST)
 DE (Thermotable phenol sulfotransferase) (Ts-PST) (HASTI/HAST2) (ST1A3).
 GN SULT1A1 OR STP1 OR STP.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=93249441; PubMed=8484775;
 RA Zhu X., Veronese M.E., Sansom L.N., McManus M.E.;
 RT "Molecular characterisation of a human aryl sulfotransferase cDNA.";
 RL Biochem. Biophys. Res. Commun. 192:671-676(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=93371391; PubMed=8363592;
 RA Zhu X., Veronese M.E., Bernard C.C., Sansom L.N., McManus M.E.;
 RT "Identification of two human brain aryl sulfotransferase cDNAs.";
 RL Biochem. Biophys. Res. Commun. 195:120-127(1993).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Hippocampus;
 RX MEDLINE=95169114; PubMed=7864863;
 RA Hwang S.-R., Kohn A.B., Hook V.Y.H.;
 RT "Molecular cloning of an isoform of phenol sulfotransferase from human brain hippocampus.";
 RL Biochem. Biophys. Res. Commun. 207:701-707(1995).
 RN [4]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Liver;
 RX MEDLINE=93140712; PubMed=8423770;
 RA Wilborn T.W., Comer K.A., Dooley T.P., Reardon I.M.,
 RA Heinrikson R.L., Falany C.N.;
 RT "Sequence analysis and expression of the cDNA for the phenol-sulfating form of human liver phenol sulfotransferase.";
 RL Mol. Pharmacol. 43:70-77(1993).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Blood;
 RX MEDLINE=95209704; PubMed=7695643;
 RA Jones A.L., Hagen M., Coughtrie M.W., Roberts R.C., Glatt H.;
 RT "Human platelet phenol sulfotransferases: cDNA cloning, stable expression in V79 cells and identification of a novel allelic variant of the phenol-sulfating form.";
 RL Biochem. Biophys. Res. Commun. 208:855-862(1995).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=96065417; PubMed=7581483;
 RA Ozawa S., Nagata K., Shimada M., Ueda M., Tazuke T., Yamazoe Y.,

RA Kato R.;
 RT "Primary structures and properties of two related forms of aryl sulfotransferases in human liver.";
 RL Pharmacogenetics 5:S135-S140(1995).
 RN [7]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94305556; PubMed=8033246;
 RA Yamazoe Y., Nagata K., Ozawa S., Kato R.;
 RT "Structural similarity and diversity of sulfotransferases.";
 RL Chem. Biol. Interact. 92:107-117(1994).
 RN [8]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96211162; PubMed=9242421;
 RA Bernier F., Soucy P., Liu-Fhe V.;
 RT "Human phenol sulfotransferase gene contains two alternative promoters: Structure and expression of the gene.";
 RL DNA Cell Biol. 15:367-375(1996).
 RN [9]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99074339; PubMed=9855620;
 RA Raftogiannis R.B., Her C., Weinshilboum R.M.;
 RT "A single amino acid, Glu46, governs the substrate specificity of human dopamine sulfotransferase, SULT1A3.";
 RL Mol. Pharmacol. 54:942-948(1998).
 RN [12]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [13]
 RP SEQUENCE OF 1-107 FROM N.A.
 RX MEDLINE=94117022; PubMed=8288252;
 RA Dooley T.P., Obermoller R.D., Leiter E.H., Chapman H.D., Falany C.N., Deng Z., Siciliano M.J.;
 RT "Mapping of the phenol sulfotransferase gene (STP) to human chromosome 16p12.1-p11.2 and to mouse chromosome 7.";
 RL Genomics 18:440-443(1993).
 RN [14]
 RP CHARACTERIZATION.
 RX MEDLINE=94379981; PubMed=8093002;
 RA Veronese M.E., Burgess W., Zhu X., McManus M.E.;
 RT "Functional characterization of two human sulphotransferase cDNAs that encode monoamine- and phenol-sulfating forms of phenol sulphotransferase: substrate kinetics, thermal-stability and inhibitor-sensitivity studies.";

Biochem. J. 302:497-502(1994).
 [15] MUTAGENESIS OF CYS-70.
 RX MEDLINE=94306582; PubMed=8033270;
 RA Falany C.N., Zhuang W., Falany J.L.;
 RT "Characterization of expressed human phenol-sulfating phenol
 RT sulfotransferase: effect of mutating cys70 on activity and
 RT thermostability.";
 RL Chem. Biol. Interact. 92:57-66(1994).
 [16]
 RN VARIANTS GLN-37; HIS-213 AND VAL-223.
 RP MEDLINE=98005125; PubMed=9345314;
 RX Raftogiannis R.B., Wood T.C., Otterness D.M., Van Loon J.A.,
 RA Weinshilboum R.M.;
 RT "Phenol sulfotransferase pharmacogenetics in humans: association of
 RT common SULT1A1 alleles with TS PST phenotype.";
 RL Biochem. Biophys. Res. Commun. 239:298-304(1997).
 [17]
 RN VARIANT HIS-213.
 RP MEDLINE=20222641; PubMed=10762004;
 RX Engelke C.E., Meinel W., Boeing H., Glatt H.;
 RA "Association between functional genetic polymorphisms of human
 RT sulfotransferases 1A1 and 1A2.";
 RL Pharmacogenetics 10:163-169(2000).
 CC -!- FUNCTION: Catalyzes the sulfate conjugation of catecholamines,
 CC phenolic drugs and neurotransmitters. Is also responsible for the
 CC sulfation and activation of minoxidil. Mediates the metabolic
 CC activation of carcinogenic N-hydroxyarylamines to DNA binding
 CC products and could so participate as modulating factor of cancer
 CC risk.
 CC -!- CATALYTIC ACTIVITY: 3'-phosphoadenylylsulfate + a phenol =
 CC adenosine 3',5'-bisphosphate + an aryl sulfate.
 CC -!- SUBUNIT: Homodimer.
 CC -!- CELLULAR LOCATION: Cytoplasmic.
 CC -!- TISSUE SPECIFICITY: Liver, lung, adrenal, brain, platelets and
 CC skin.
 CC -!- SIMILARITY: Belongs to the sulfotransferase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U09031; AAA18613.1; -
 CC EMBL; L19955; AAA02935.1; -
 CC EMBL; L10819; AAA35562.1; -
 CC EMBL; U26309; AAA67895.1; -
 CC EMBL; L19959; AAA99892.1; -
 CC EMBL; X84654; CAA59147.1; -
 CC EMBL; X78283; CAA55089.1; -
 CC EMBL; U54701; AAC50480.1; -
 CC EMBL; L15346; AAA60595.1; -
 CC EMBL; U52852; AAC51816.1; -
 CC EMBL; AJ007418; CAA07495.1; -
 CC EMBL; EC000923; AAH00923.1; -
 CC EMBL; U71086; AAB09597.1; -
 CC PIR; I57945; I57945.
 CC PIR; JC2523; JC2523.
 CC PIR; JC5248; JC5248.
 CC PIR; S52794; S52794.
 CC HSP; F50224; ICUM.
 CC DR Genew; HGNC:11453; SULT1A1.
 CC MIM; 171150; -
 CC GO; GO:0008146; F:sulfotransferase activity; TAS.
 CC GO; GO:0009308; P:amine metabolism; TAS.
 CC InterPro; IPR000863; Sulfotransferase.
 CC Pfam; PF00685; Sulfotransfer; 1.

Query Match

Best Local Similarity 56.9%; Score 37; DB 1; Length 295;
Pred. No. 34;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 3 WDSGTHPV 11
 Db 180 WELSRTHPV 188
 RESULT 13
 SUP2_HUMAN
 ID SUP2_HUMAN STANDARD; PRT; 295 AA.
 AC P50226; P78393;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Phenol-sulfating phenol sulfotransferase 2 (EC 2.8.2.1) (P-PST)
 DE (STIA2).
 GN SULT1A2 OR STP2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Zhu X., Veronese M.E., Iocco P., McManus M.E.;
 RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=96065417; PubMed=7581483;
 RA Ozawa S., Nagata K., Shimada M., Ueda M., Tezuka T., Yamazoe Y.,
 RA Kato R.;
 RT "Primary structures and properties of two related forms of aryl
 RT sulfotransferases in human liver.";
 RL Pharmacogenetics 5:S135-S140(1995).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94306556; PubMed=8033246;
 RA Yamazoe Y., Nagata K., Ozawa S., Kato R.;
 RT "Structural similarity and diversity of sulfotransferases.";
 RL Chem. Biol. Interact. 92:107-117(1994).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96299636; PubMed=8661000;
 RA Her C., Raftogiannis R., Weinshilboum R.M.;
 RT "Human phenol sulfotransferase STP2 gene: molecular cloning,
 RT structural characterization, and chromosomal localization.";
 RL Genomics 33:409-420(1996).
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Dooley T.P., Huang Z.;
 RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97237042; PubMed=9119390;
 RA Gaedigk A., Beatty B.G., Grant D.M.;
 RT "Cloning, structural organization, and chromosomal mapping of the
 RT human phenol sulfotransferase STP2 gene.";
 RL Genomics 40:242-246(1997).
 RN [7]
 RP CHARACTERIZATION.
 RX MEDLINE=95196680; PubMed=7889867;
 RA Yamazoe Y., Ozawa S., Nagata K., Cong D.-W., Kato R.;
 RT "Characterization and expression of hepatic sulfotransferase involved
 RT in the metabolism of N-substituted aryl compounds.";
 RL Environ. Health Perspect. 102:99-103(1994).
 RN [8]
 RP VARIANT ASN-235.
 RX MEDLINE=20222641; PubMed=10762004;
 RA Engelke C.E., Meinel W., Boeing H., Glatt H.;
 RT "Association between functional genetic polymorphisms of human
 RT sulfotransferases 1A1 and 1A2.";
 RL Pharmacogenetics 10:163-169(2000).

CC -!- FUNCTION: Catalyzes the sulfate conjugation of catecholamines,
CC phenolic drugs and neurotransmitters. Is also responsible for the
CC sulfation and activation of minoxidil. Mediates the metabolic
CC activation of carcinogenic N-hydroxyarylamines to DNA binding
CC products and could so participate as modulating factor of cancer
CC risk.
CC
CC -!- CATALYTIC ACTIVITY: 3'-phosphoadenylylsulfate + a phenol =
CC adenosine 3',5'-bisphosphate + an aryl sulfate.
CC
CC -!- SUBUNIT: Homodimer (By similarity).
CC
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC
CC -!- SIMILARITY: Belongs to the sulfotransferase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; U28170; AAB09659.1; -
CC DR EMBL; U28169; AAB09658.1; -
CC DR EMBL; X78282; CA255088.1; -
CC DR EMBL; U34804; AAB09758.1; -
CC DR EMBL; U72202; AAB08970.1; -
CC DR EMBL; U72196; AAB08970.1; JOINED.
CC DR EMBL; U72197; AAB08970.1; JOINED.
CC DR EMBL; U72198; AAB08970.1; JOINED.
CC DR EMBL; U72199; AAB08970.1; JOINED.
CC DR EMBL; U72200; AAB08970.1; JOINED.
CC DR EMBL; U72201; AAB08970.1; JOINED.
CC DR EMBL; U76619; AAB18753.1; -
CC DR EMBL; U33886; AAC51149.1; -
CC DR PIR; G01843; G01843.
CC DR PIR; JC5249; JC5249.
CC DR HSP; P50224; ICJM.
CC DR Genew; HGNC:11454; SULT1A2.
CC MIM; 601292; -
CC DR GO; GO:0008146; P:sulfotransferase activity; TAS.
CC DR GO; GO:0009309; P:amine biosynthesis; TAS.
CC InterPro; IPR000863; Sulfotransferase.
CC Pfam; PF00685; Sulfotransferase; 1.
CC ProDom; PD001218; Sulfotransferase; 1.
CC Catecholamine metabolism; Steroid metabolism; Transferase;
KW Polymorphism.
FT BINDING 259 265 PAPs BINDING SITE (POTENTIAL).
FT VARIANT 7 7 T -> I
FT FTID=VAR_007426.
FT VARIANT 235 235 T -> N.
FT FTID=VAR_007427.
FT CONFLICT 19 19 P -> L (IN REF. 2).
FT CONFLICT 290 290 S -> T (IN REF. 6).
SQ SEQUENCE 295 AA; 34285 MW; FAD68E34849BE5A CRC64;

Query Match 56.9%; Score 37; DB 1; Length 295;
Best Local Similarity 66.7%; Pred. No. 34;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 WDSSGTHPV 11
DB 180 WELSRTHPV 188

RESULT 14
SUPM_HUMAN
ID SUPM_HUMAN STANDARD; PRT; 295 AA.
AC P50224;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Monamine-sulfating phenol sulfotransferase (EC 2.8.2.1)
DE (Sulfotransferase, monamine-prefering) (M-PST) (thermolabile phenol
DE sulfotransferase) (TL-PST) (Placental estrogen sulfotransferase)

DE (Catecholamine-sulfating phenol sulfotransferase) (HAST3).
GN SULT1A3 OR STM.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBT_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=93371391; PubMed=8363592;
RA Zhu X., Veronese M.E., Bernard C.C., Sansom L.N., McManus M.E.;
RT "Identification of two human brain aryl sulfotransferase cDNAs.";
RL Biochem. Biophys. Res. Commun. 195:120-127(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=94244843; PubMed=8187949;
RA Bernier F., Lopez-Solache I., Labrie F., Luu-The V.;
RT "Cloning and expression of cDNA encoding human placental estrogen
RT sulfotransferase.";
RL Mol. Cell. Endocrinol. 99:R11-R15(1994).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=94161723; PubMed=8117269;
RA Wood T.C., Aksoy I.A., Aksoy S., Weinshilboum R.M.;
RT "Human liver thermolabile phenol sulfotransferase: cDNA cloning,
RT expression and characterization.";
RL Biochem. Biophys. Res. Commun. 198:1119-1127(1994).
RN [4]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 84-101.
RC TISSUE=Liver;
RX MEDLINE=95100963; PubMed=7802665;
RA Dooley T.P., Probst P., Munroe P.B., Mole S.E., Liu Z., Doggett N.A.;
RT "Genomic organization and DNA sequence of the human catecholamine-
RT sulfating phenol sulfotransferase gene (STM).";
RL Biochem. Biophys. Res. Commun. 205:1325-1332(1994).
RN [5]
RP SEQUENCE FROM N.A.
RC TISSUE=Platelet;
RX MEDLINE=95209696; PubMed=7695637;
RA Aksoy I.A., Weinshilboum R.M.;
RT "Human thermolabile phenol sulfotransferase gene (STM): molecular
RT cloning and structural characterization.";
RL Biochem. Biophys. Res. Commun. 208:786-795(1995).
RN [6]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=95209704; PubMed=7695643;
RA Jones A.L., Hagen M., Coughtrie M.W.H., Roberts R.C., Glatt H.;
RT "Human platelet phenol sulfotransferases: cDNA cloning, stable
RT expression in V79 cells and identification of a novel allelic variant
RT of the phenol-sulfating form.";
RL Biochem. Biophys. Res. Commun. 208:855-862(1995).
RN [7]
RP SEQUENCE FROM N.A.
RC TISSUE=Leukocyte;
RX MEDLINE=95050600; PubMed=7961757;
RA Bernier F., Leblanc G., Labrie F., Luu-The V.;
RT "Structure of human estrogen and aryl sulfotransferase gene. Two mRNA
RT species issued from a single gene.";
RL J. Biol. Chem. 269:28200-28205(1994).
RN [8]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ugin T.B., Toehiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettunen M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson W.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [9]
RP SEQUENCE OF 139-198 FROM N.A.
RC TISSUE=Lymphocytes;
RX MEDLINE=95130098; PubMed=7829089;
RA Aksoy I.A., Callen D.F., Apostolou S., Her C., Weinshilboum R.M.;
RT "Thermolabile phenol sulfotransferase gene (STM): localization to
RT human chromosome 16p11.2.";
RL Genomics 23:275-277(1994).
RN [10]
RP CHARACTERIZATION.
RX MEDLINE=94379981; PubMed=8093002;
RA Veronese M.E., Burgess W., Zhu X., McManus M.E.;
RT "Functional characterization of two human sulfotransferase cDNAs
RT that encode monamine- and phenol-sulphating forms of phenol
RT sulfotransferase: substrate kinetics, thermal-stability and
RT inhibitor-sensitivity studies.";
RL Biochem. J. 302:497-502(1994).
RN [11]
RP X-RAY CRYSTALLOGRAPHY (2.4 ÅNGSTROMS).
RC TISSUE=Brain;
RX MEDLINE=20013065; PubMed=10543947;
RA Bidwell L.M., McManus M.E., Gaedigk A., Kakuta Y., Negishi M.,
RA Pedersen L., Martin J.L.;
RT "Crystal structure of human catecholamine sulfotransferase.";
RL J. Mol. Biol. 293:521-530(1999).
CC -!- FUNCTION: Catalyzes the sulfate conjugation of phenolic monoamines
CC (neurotransmitters such as dopamine, norepinephrine and serotonin)
CC and phenolic and catechol drugs.
CC -!- CATALYTIC ACTIVITY: 3'-phosphoadenylylsulfate + a phenol =
CC adenosine 3',5'-bisphosphate + an aryl sulfate.
CC -!- SUBUNIT: Homodimer.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- TISSUE SPECIFICITY: Liver, colon, kidney, lung, brain, spleen,
CC small intestine, placenta and Leukocyte.
CC -!- PTM: The N-terminus is blocked.
CC -!- SIMILARITY: Belongs to the sulfotransferase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L19956; AAA02943.1; -;
DR EMBL; L25275; AAA36523.1; -;
DR EMBL; U08032; AAA17723.1; -;
DR EMBL; U20499; AAA64490.1; -;
DR EMBL; X84653; CAA59146.1; -;
DR EMBL; L34160; -; NOT ANNOTATED_CDS.
DR EMBL; U37686; AAA86536.1; -;
DR EMBL; BC014471; AAA14471.1; -;
DR EMBL; U08099; AAA82126.1; -;
DR PIR; A55451; A55451.
DR PDB; 1CJM; 10-NOV-99.
DR Genew; HGNC:11455; SULT1A3.
DR MIM; 600641; -;
DR GO; GO:0005737; Cytoplasm; TAS.
DR GO; GO:0004062; F:aryl sulfotransferase activity; TAS.
DR InterPro; IPR000863; Sulfotransferase.

DR Pfam; PF00685; Sulfotransferase; 1.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Catecholamine metabolism; Steroid metabolism; Transferase;
KW 3D-structure.
FT BINDING 259 265 PAPS BINDING SITE (POTENTIAL).
FT STRAND 12 15
FT TURN 16 17
FT STRAND 18 21
FT HELIX 22 27
FT TURN 28 28
FT HELIX 30 33
FT TURN 38 39
FT STRAND 41 45
FT TURN 47 48
FT HELIX 51 62
FT STRAND 81 81
FT TURN 82 83
FT TURN 87 88
FT TURN 95 98
FT STRAND 104 107
FT HELIX 111 113
FT HELIX 116 120
FT TURN 121 122
FT STRAND 124 129
FT HELIX 132 145
FT TURN 147 148
FT HELIX 155 163
FT TURN 164 165
FT TURN 168 169
FT HELIX 172 182
FT TURN 183 185
FT STRAND 188 192
FT HELIX 193 198
FT TURN 200 211
FT TURN 212 212
FT HELIX 264 266
FT TURN 267 267
FT TURN 270 283
FT TURN 284 284
SQ SEQUENCE 295 AA; 34196 MW; ECDDBC03DBE30D46 CRC64;
Query Match 56.9%; Score 37; DB 1; Length 295;
Best Local Similarity 66.7%; Pred. No. 34;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Oy 3 WDSGTHPV 11
|:|
Db 180 WELSRTHPV 188
RESULT 15
SUPP_MACFA
ID SUPP_MACFA STANDARD; PRT; 295 AA.
AC P52846;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Phenol-sulfating phenol sulfotransferase (EC 2.8.2.1) (P-PST).
GN STP.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9541;
RN [1].
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Ogura K., Satsukawa M., Okuda H., Watabe T.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Catalyzes the sulfate conjugation of catecholamines and
CC phenolic drugs and neurotransmitters (By similarity).
CC -!- CATALYTIC ACTIVITY: 3'-phosphoadenylylsulfate + a phenol =
CC adenosine 3',5'-bisphosphate + an aryl sulfate.

CC -!- SUBUNIT: Homodimer (By similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- SIMILARITY: Belongs to the sulfotransferase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; D85514; BAA12822.1; -;
 DR PIR; G02924; G02924.
 DR HSP; P50224; 1CJM.
 DR InterPro; IPR000863; Sulfotransferase.
 DR Pfam; PF00685; Sulfotransfer; 1.
 DR ProDom; PD001218; Sulfotransferase; 1.
 KW Catecholamine metabolism; Steroid metabolism; Transferase.
 FT BINDING 259 265 PAPS BINDING SITE (POTENTIAL).
 SQ SEQUENCE 295 AA; 34238 MW; BD51639D1570A841 CRC64;

Query Match 56.9%; Score 37; DB 1; Length 295;
 Best Local Similarity 66.7%; Pred. No. 34;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 WDSSGTHPV 11
 | : | | | |
 Db 180 WELSHTEPV 188

RESULT 16
 EXL2 HUMAN STANDARD; PRT; 330 AA.
 AC Q9UB06;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Exostosin-like 2 (EC 2.4.1.223) (Glucuronyl-galactosyl-proteoglycan 4-
 DE alpha-N-acetylglucosaminyltransferase) (Alpha-1,4-N-
 DE acetylhexosaminyltransferase EXTL2) (Alpha-GalNAc EXTL2) (EXT-related
 DE protein 2).
 GN EXTL2 OR EXTR2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98111853; PubMed=9450183;
 RA Wuyts W., Van Hul W., Hendrickx J., Speleman F., Wauters J.,
 RA De Bouille K., Van Roy N., Van Agtmael T., Bossuyt P., Willems P.J.;
 RT "Identification and characterization of a novel member of the EXT gene
 RT family, EXTL2.";
 RL Eur. J. Hum. Genet. 5:382-389(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98139867; PubMed=9473480;
 RA Saito T., Seki N., Yamauchi M., Tsuji S., Hayaishi A., Kozuma S.,
 RA Hori T.-A.;
 RT "Structure, chromosomal location, and expression profile of EXTR1 and
 RT EXTR2, new members of the multiple exostosins gene family.";
 RL Biochem. Biophys. Res. Commun. 243:61-66(1998).
 RN [3]
 RP SUBCELLULAR LOCATION.
 RX MEDLINE=20105530; PubMed=10639137;
 RA McCormick C., Duncan G., Goutsos K.T., Tufaro F.;
 RT "The putative tumor suppressors EXTR1 and EXTR2 form a stable complex
 RT that accumulates in the Golgi apparatus and catalyzes the synthesis
 RT of heparan sulfate.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:668-673(2000).
 RN [4]
 RP FUNCTION, AND SEQUENCE OF 54-83.

RX MEDLINE=99253945; PubMed=10318803;
 RA Kitagawa H., Shimakawa H., Sugahara K.;
 RT "The tumor suppressor EXT-like gene EXTL2 encodes an alpha1,
 RT 4-N-acetylhexosaminyltransferase that transfers N-acetylgalactosamine
 RT and N-acetylglucosamine to the common glycosaminoglycan-protein
 RT linkage region. The key enzyme for the chain initiation of heparan
 RT sulfate.";
 RL J. Biol. Chem. 274:13933-13937(1999).
 CC -!- FUNCTION: Glycosyltransferase required for the biosynthesis of
 CC heparan-sulfate and responsible for the alternating addition of
 CC beta-1,4-linked glucuronic acid (GlcA) and alpha-1,4-linked N-
 CC acetylglucosamine (GlcNAc) units to nascent heparan sulfate
 CC chains.
 CC -!- CATALYTIC ACTIVITY: UDP-N-acetyl-D-glucosamine + beta-D-
 CC glucuronosyl-(1->3)-beta-D-galactosyl-(1->3)-beta-D-galactosyl-
 CC (1->4)-beta-D-xylosyl-proteoglycan = UDP + alpha-N-acetyl-D-
 CC glucosaminyl-(1->4)-beta-D-glucuronosyl-(1->3)-beta-D-galactosyl-
 CC (1->3)-beta-D-galactosyl-(1->4)-beta-D-xylosyl-proteoglycan.
 CC -!- PATHWAY: Heparin/heparan sulfate synthesis.
 CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Endoplasmic
 CC reticulum (By similarity). A soluble form is found in the serum.
 CC -!- TISSUE SPECIFICITY: Ubiquitous.
 CC -!- PTM: The soluble form derives from the membrane form by
 CC proteolytic processing.
 CC -!- SIMILARITY: Belongs to the glycosyltransferase family 47.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF000416; AAC02898.1; -;
 DR EMBL; AB009284; BAA24081.1; -;
 DR PIR; JC5935; JC5935.
 DR Genew; HGNC:3516; EXTL2.
 DR MIM; 602411; -;
 KW Transferase; Glycosyltransferase; Endoplasmic reticulum;
 KW Transmembrane; Signal-anchor; Glycoprotein.
 FT DOMAIN 1 22 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 23 43 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
 FT FT
 FT DOMAIN 44 330 LUMENAL (POTENTIAL).
 FT SITE 53 54 CLEAVAGE.
 FT CATHID 74 74 N-LINKED (GLCNAC. .).
 SQ SEQUENCE 330 AA; 37465 MW; 6976BE7EC6F588C8 CRC64;

Query Match 56.9%; Score 37; DB 1; Length 330;
 Best Local Similarity 55.6%; Pred. No. 39;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 3 WDSSGTHPV 11
 | : | | | |
 Db 111 WNSLGFHFI 119

RESULT 17
 EXL2 MOUSE STANDARD; PRT; 330 AA.
 AC Q9ES89; Q9CX90;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Exostosin-like 2 (EC 2.4.1.223) (Glucuronyl-galactosyl-proteoglycan 4-
 DE alpha-N-acetylglucosaminyltransferase) (Alpha-1,4-N-
 DE acetylhexosaminyltransferase EXTL2) (Alpha-GalNAc EXTL2) (EXT-related
 DE protein 2).
 GN EXTL2 OR EXTR2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC

OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20422478; PubMed=10965119;
RA Wuyts W., Van Hul W.;
RT "Characterization and genomic localization of the mouse Extl2 gene.";
RL Cytogenet. Cell Genet. 89:185-188 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryonic head;
RX MEDLINE=22354683; PubMed=12466851;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RN Nikaide I., Oseato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuoka H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Bruscia V., Chochia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramchandran S.,
RA Ravasi T., Reed J.C., Reid D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai K., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume M., Imorani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yabuuchi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RN "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Mammary gland;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bobak S.A., McWeeney J.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S.J., Garcia A.M., Gay L.J., Rulyk S.W.,
RA Villalon D.K., Muzny D.K., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywicki M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RN "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -!- FUNCTION: Glycosyltransferase required for the biosynthesis of
CC heparan-sulfate and responsible for the alternating addition of
CC beta-1-4-linked glucuronic acid (GlcA) and alpha-1-4-linked N-
CC acetylglucosamine (GlcNAc) units to nascent heparan sulfate
CC chains (By similarity).
CC -!- CATALYTIC ACTIVITY: UDP-N-acetyl-D-glucosamine + beta-D-
CC glucuronosyl-(1->3)-beta-D-galactosyl-(1->3)-beta-D-galactosyl-
CC (1->4)-beta-D-xylosyl-proteoglycan = UDP + alpha-N-acetyl-D-
CC glucosaminyl-(1->4)-beta-D-glucuronosyl-(1->3)-beta-D-galactosyl-
CC

(1->3)-beta-D-galactosyl-(1->4)-beta-D-xylosyl-proteoglycan.
-!- PATHWAY: Heparin/heparan sulfate synthesis.
-!- SUBCELLULAR LOCATION: Type II membrane protein. Endoplasmic
reticulum (By similarity).
-!- SIMILARITY: Belongs to the glycosyltransferase family 47.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
EMBL; AF200973; AAG17542.1; -
EMBL; AK019370; BAB31683.2; -
EMBL; BC031438; AAH31438.1; -
MGD; MGI:1889574; Extl2.
Transferase; Glycosyltransferase; Endoplasmic reticulum;
Transmembrane; Signal-anchor; Glycoprotein.
DOMAIN 1 21 CYTOPLASMIC (POTENTIAL).
TRANSMEM 22 42 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
(POTENTIAL).
DOMAIN 43 330 LUMENAL (POTENTIAL).
FT CARBOHYD 75 75 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 191 191 Y -> D (IN REF. 2).
SQ SEQUENCE 330 AA; 37390 MW; 51F88B5D3E5EADF CRC64;
Query Match 56.9%; Score 37; DB 1; Length 330;
Best Local Similarity 55.6%; Pred. No. 39;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Oy 3 WDSSTGTPV 11
Db 112 WNSLGPPI 120
RESULT 18
PELB ERWCH STANDARD; PRT; 375 AA.
AC PE4959;
DT 13-AUG-1997 (Rel. 05, Created)
DT 13-AUG-1997 (Rel. 05, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Pectate lyase B precursor (EC 4.2.2.2).
GN PELB.
OS Erwinia chrysanthemi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pectobacterium.
OX NCBI_TaxID=556;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EC16;
RX MEDLINE=87056939; PubMed=3536853;
RA Keen N.T., Tamaki S.;
RT "Structure of two pectate lyase genes from Erwinia chrysanthemi EC16
and their high-level expression in Escherichia coli.";
RL J. Bacteriol. 168:595-606 (1986).
CC -!- FUNCTION: Involved in maceration and soft-rotting of plant tissue.
CC -!- CATALYTIC ACTIVITY: Eliminative cleavage of (1->4)-alpha-D-
CC galacturonan to give oligosaccharides with 4-deoxy-alpha-D-galact-
CC 4-enuronosyl groups at their non-reducing ends.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the polysaccharide lyase family 1. PLADES
CC subfamily.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

```

CC -----
DR EMBL; M14510; AAA24847.1; -.
DR PIR; B25158; WZWCPCB.
DR HSPG; P11073; 1AIR.
DR InterPro; IPR002022; Amb_allergen.
DR Pfam; PF00544; pec_lyase; 1.
DR SMART; SM00656; Amb_all; 1.
DR Lyase; Multigene family; Signal.
KW SIGNAL 1 22
FT CHAIN 23 375 PECTATE LYASE B.
FT DISULFID 93 176 BY SIMILARITY.
FT DISULFID 351 374 BY SIMILARITY.
FT ACT_SITE 240 240 POTENTIAL.
SQ SEQUENCE 375 AA; 40234 MW; F3FE75D507B8F883 CRC64;

Query Match 56.9%; Score 37; DB 1; Length 375;
Best Local Similarity 50.0%; Pred No. 45;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 NSWDSSGTHP 10
   :|||:|:|:|
Db 329 DTWTSTGYTP 338

RESULT 19
PELC_ERWCH STANDARD; PRT; 375 AA.
ID PELC_ERWCH STANDARD; PRT; 375 AA.
AC P11073;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Pectate lyase C precursor (EC 4.2.2.2).
GN PELC.
OS Erwinia chrysanthemi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pectobacterium.
OX NCBI_TaxID=556;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EC16;
RX MEDLINE=88298652; PubMed=3042750;
RA Tamaki S.J., Gold S., Robeson M., Manulis S., Keen N.T.;
RT "Structure and organization of the pel genes from Erwinia
   chrysanthemi EC16.";
RL J. Bacteriol. 170:3468-3478 (1988).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
RX MEDLINE=93276270; PubMed=8502994;
RA Yoder M.D., Keen N.T., Jurnak F.;
RT "New domain motif: the structure of pectate lyase C, a secreted plant
   virulence factor.";
RL Science 260:1503-1507 (1993).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
RA Lietzke S.E., Scavetta R.D., Yoder M.D., Jurnak F.A.;
RT "The refined three-dimensional structure of pectate lyase E from
   Erwinia chrysanthemi at 2.2-A resolution.";
RL Plant Physiol. 111:73-92 (1996).
CC 1- FUNCTION: Involved in maceration and soft-rotting of plant tissue.
CC 4-enuronosyl groups at their non-reducing ends.
CC 1- CATALYTIC ACTIVITY: Eliminative cleavage of (1->4)-alpha-D-
   galacturonan to give oligosaccharides with 4-deoxy-alpha-D-galact-
   4-enuronosyl groups at their non-reducing ends.
CC 1- SUBCELLULAR LOCATION: Secreted.
CC 1- SIMILARITY: Belongs to the polysaccharide lyase family 1. PLADES
   subfamily.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC -----
DR EMBL; M19411; AAA24849.1; -.
DR PIR; A31091; WZWC6C.
DR PDB; 2PEC; 14-FEB-95.
DR PDB; 1AIR; 16-JUN-97.
DR PDB; 1O8H; 03-APR-03.
DR PDB; 1O8J; 03-APR-03.
DR PDB; 1O8D; 03-APR-03.
DR PDB; 1O8E; 03-APR-03.
DR PDB; 1O8F; 03-APR-03.
DR PDB; 1O8H; 03-APR-03.
DR PDB; 1O8I; 03-APR-03.
DR PDB; 1O8J; 03-APR-03.
DR PDB; 1O8K; 03-APR-03.
DR PDB; 1O8L; 03-APR-03.
DR PDB; 1O8M; 03-APR-03.
DR PDB; 1PLU; 13-JUL-99.
DR InterPro; IPR002022; Amb_allergen.
DR Pfam; PF00544; pec_lyase; 1.
DR SMART; SM00656; Amb_all; 1.
DR Lyase; Multigene family; Signal; 3D-structure.
KW SIGNAL 1 22
FT CHAIN 23 375 PECTATE LYASE C.
FT DISULFID 94 177
FT DISULFID 351 374
FT ACT_SITE 240 240
FT STRAND 25 25
FT STRAND 31 31
FT TURN 34 35
FT TURN 38 39
FT STRAND 41 45
FT HELIX 48 57
FT TURN 58 59
FT STRAND 60 60
FT TURN 62 63
FT STRAND 66 66
FT TURN 68 68
FT STRAND 69 69
FT STRAND 74 78
FT HELIX 83 90
FT TURN 91 91
FT HELIX 93 95
FT TURN 97 98
FT STRAND 103 107
FT STRAND 109 109
FT STRAND 113 117
FT TURN 119 120
FT STRAND 123 123
FT STRAND 126 130
FT STRAND 135 138
FT STRAND 141 143
FT HELIX 148 150
FT TURN 151 151
FT STRAND 155 158
FT STRAND 163 166
FT STRAND 169 171
FT TURN 178 179
FT HELIX 181 183
FT STRAND 191 194
FT STRAND 199 203
FT STRAND 206 212
FT STRAND 215 216
FT TURN 220 221
FT STRAND 226 230
FT STRAND 233 239
FT STRAND 243 245
FT STRAND 248 252
FT STRAND 255 259
FT STRAND 264 267
FT TURN 268 269
FT STRAND 271 275
FT STRAND 278 283
FT STRAND 286 288
FT STRAND 297 301

```

FT TURN 308 308
 FT HELIX 309 314
 FT TURN 315 315
 FT STRAND 316 317
 FT STRAND 326 327
 FT TURN 329 330
 FT HELIX 349 355
 FT HELIX 356 358
 FT TURN 359 359
 FT TURN 365 366
 FT HELIX 371 373
 SQ SEQUENCE 375 AA; 39943 MW; F76DD8195A35B886 CRC64;

Query Match
 Best Local Similarity 56.9%; Score 37; DB 1; Length 375;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NSWSSSGTHP 10
 :|||:|
 Db 329 DSWTSTGTFF 338

RESULT 20
 ENV_MLVHO STANDARD; PRT; 666 AA.
 AC P21436;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-MAY-1991 (Rel. 18, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE ENV polyprotein precursor [Contains: Knob protein GP76; Spike protein P15E].
 GN ENV.
 OS HomoVirus murine leukemia virus (Mus hortulanus virus).
 OC Viruses; Retroviral viruses; Retroviridae; Gammaretrovirus.
 OX NCBI_TaxID=11799;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90051094; PubMed=2554579;
 RA Vaytek P., Kozak C.A.;
 RT "Nucleotide sequence and mode of transmission of the wild mouse
 ecotropic virus, HoMuLV.";
 RL Virology 173:58-67(1989).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M26527; -; NOT_ANNOTATED_CDS.
 DR PIR; B32594; VCMVHL.
 DR HSP; P03385; IMOF.
 DR InterPro; IPR002050; Env polyprotein.
 DR InterPro; IPR008981; F MuLV bind.
 DR Pfam; PF00429; Env polyprotein; 1.
 KW Coat protein; Glycoprotein; Polyprotein; Signal; Transmembrane.
 FT SIGNAL 1 30
 FT CHAIN 31 466
 FT CHAIN 467 666
 FT TRANSMEM 473 489
 FT TRANSMEM 607 623
 FT CARBOHYD 42 42
 FT CARBOHYD 197 197
 FT CARBOHYD 290 290
 FT CARBOHYD 324 324
 FT CARBOHYD 356 356
 FT CARBOHYD 363 363
 FT CARBOHYD 431 431
 SQ SEQUENCE 666 AA; 73034 MW; FDC779564B213D1 CRC64;

Query Match
 Best Local Similarity 56.9%; Score 37; DB 1; Length 666;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Best Local Similarity 55.6%; Pred. No. 83;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 WDSGTHPV 11
 :|||:|
 Db 56 WATSGSHPL 64

RESULT 21
 GPRS DROME
 ID GPRS DROME STANDARD; PRT; 1302 AA.
 AC O61366; Q9V7S2;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Serine-enriched protein.
 GN GPRS OR C618471/CG5967.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt J.G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos F.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieri S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Godek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laoko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J.J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Stadling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 RN [2].
 RP REVISIONS.
 RX MEDLINE=22426069; PubMed=12537572;
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo A.J., Schroeder A.J., Shu S.O.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,

```
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review.";
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
[3]
RP SEQUENCE OF 526-1302 FROM N.A.
RC STRAIN=Canton-S;
RA Da Lage J.-L.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Contains 1 BTB/POZ domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
EMBL; AF022713; AAD09150.2; -.
DR FlyBase; FBgn0024232; GBPs.
DR InterPro; IPR000210; BTB_POZ.
DR SMART; SM00225; BTB; 1.
DR PROSITE; PS50097; BTB; 1.
DR DOMAIN 40 158
FT CONFLICT 691 691 E -> D (IN REF. 3).
FT CONFLICT 744 744 Q -> H (IN REF. 3).
FT CONFLICT 788 788 D -> V (IN REF. 3).
FT CONFLICT 808 808 Q -> L (IN REF. 3).
FT CONFLICT 894 894 D -> V (IN REF. 3).
FT CONFLICT 1167 1167 R -> G (IN REF. 3).
FT CONFLICT 1301 1301 S -> N (IN REF. 3).
SQ SEQUENCE 1302 AA; 144768 MW; DA3603B725140215 CRC64;

Query Match 56.9%; Score 37; DB 1; Length 1302;
Best Local Similarity 60.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSWSSGTHP 10
DB 1121 DSYDSGSGYP 1130
:|||||:
:|||||:

RESULT 22
YV04 MYCTU STANDARD; PRT; 234 AA.
AC Q50721.
DT 01-NOV-1997 (Rel. 35, Created)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein RV3404C/MT3512/MB3438C precursor.
GN RV3404C OR MT3512 OR MTCV78.24 OR MB3438C.
OS Mycobacterium tuberculosis, and
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773, 1765;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=M.tuberculosis; STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230;
RA Gordon S.V., Biglmeier K., Gas S., Barry C.E. III, Tekai F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jegels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares R.,
RA Sultston J.E., Taylor K., Whitehead S., Barrall B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).

[2]
RP SEQUENCE FROM N.A.
RC SPECIES=M.tuberculosis; STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., Deboy R., Dodson R., Gwinn M., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
RA Decher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL J. Bacteriol. 184:5479-5490(2002).
[3]
RP SEQUENCE FROM N.A.
RC SPECIES=M.bovis; STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12788972;
RA Garnier T., Biglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrall B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
CC -!- SIMILARITY: SOME, TO METHIONYL-TRNA FORMYLTRANSFERASE
CC (EC 2.1.2.9).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
EMBL; Z77165; CAB01019.1; -.
DR EMBL; AE007157; AAK47850.1; -.
DR EMBL; BX248346; CAD95625.1; -.
DR PIR; A70736; A70736.
DR TIGR; MT3512; -.
DR TuberculList; RV3404C; -.
DR InterPro; IPR002376; formyl_transf.
DR Pfam; PF00551; formyl_transf; 1.
KW Hypothetical protein; _Signal; Complete proteome.
FT SIGNAL 1 16 POTENTIAL.
FT CHAIN 17 234 HYPOTHETICAL PROTEIN
FT RV3404C/MT3512/MB3438C.
SQ SEQUENCE 234 AA; 26515 MW; 63FF857BB6FFAA8F CRC64;

Query Match 55.4%; Score 36; DB 1; Length 234;
Best Local Similarity 85.7%; Pred. No. 40;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 SWDSSGT 8
DB 130 SWDSSGS 136
|||||:
|||||:

RESULT 23
AGL9 ARADE STANDARD; PRT; 250 AA.
AC Q38694.
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Agamou-like MADS box protein AGL9 homolog (OM1).
OS Aranda deborah (Orchid).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Orchidaceae;
OC Epidendroideae; higher Epidendroideae; Vandaeae; Aeridinae;
OC Arachnis x Vanda.
OX NCBI_TaxID=29714;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94072738; PubMed=8251643;
```

RA Lu 2.X., Wu M., Loh C.S., Yeong C.Y., Goh C.J.;
 RT "Nucleotide sequence of a flower-specific MADS box cDNA clone from
 orchid."
 RL Plant Mol. Biol. 23:901-904(1993).
 CC -!- FUNCTION: Probable transcription factor active in inflorescence,
 CC development and floral organogenesis (By similarity).
 CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
 CC -!- TISSUE SPECIFICITY: Expressed in petals and weakly in sepals but
 CC not in the column (gynostemium).
 CC -!- DEVELOPMENTAL STAGE: Expressed in mature flowers and not in young
 CC developing inflorescences or young floral buds.
 CC -!- SIMILARITY: Contains 1 K-box dimerization domain.
 CC -!- SIMILARITY: Contains 1 MADS-box domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X69107; CAA48959.1; -
 DR PIR; S40405; S40405.
 DR HSP; P11746; 1MN.
 DR TRANSFAC; T03114; -
 DR InterPro; IPR002487; TF_Kbox.
 DR InterPro; IPR002100; TF_MADSbox.
 DR Pfam; PF01486; K-box; 1.
 DR Pfam; PF00319; SRP-TF; 1.
 DR SMART; SM00432; MADS; 1.
 DR PROSITE; PS00350; MADS_BOX_1; 1.
 DR PROSITE; PS00066; MADS_BOX_2; 1.
 DR PROSITE; PS00066; MADS_BOX_2; 1.
 KW Transcription regulation; DNA-binding; Nuclear protein.
 FT DOMAIN 3 57 MADS-box.
 FT DOMAIN 96 168 K-BOX.
 SQ SEQUENCE 250 AA; 28770 MW; 01EF94DADC499C41 CRC64;
 Query Match 55.4%; Score 36; DB 1; Length 250;
 Best Local Similarity 66.7%; Pred. No. 43;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 3 WDSSGTHPV 11
 |||||
 DB 163 WDPSNTHAV 191
 RESULT 24
 ID PI52 HUMAN STANDARD; PRT; 406 AA.
 AC P48426.
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Phosphatidylinositol-4-phosphate 5-kinase type II alpha (EC 2.7.1.149)
 DE (PIP5KII-alpha) (1-phosphatidylinositol-4-phosphate 5-kinase)
 DE (PtdIns(4)P-5-kinase B isoform) (Diphosphoinositide kinase).
 GN PIP5K2A OR PIP5K2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Placenta;
 RX MEDLINE=95155363; PubMed=7852364;
 RA Boronenkov I.V., Anderson R.A.;
 RT "The sequence of phosphatidylinositol-4-phosphate 5-kinase defines a
 RT novel family of lipid kinases."
 RL J. Biol. Chem. 270:2881-2884(1995).
 RN [2]
 RP REVISIONS TO 298-310 AND 381-382.

RA Boronenkov I.V.;
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Hippocampus;
 RX MEDLINE=22389257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosa A.S., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -!- FUNCTION: Catalyzes the phosphorylation of phosphatidylinositol-4-
 CC phosphate on the fifth hydroxyl of the myo-inositol ring, to form
 CC phosphatidylinositol-4,5-bisphosphate.
 CC -!- CATALYTIC ACTIVITY: ATP + 1-phosphatidyl-1D-myo-inositol 4-
 CC monophosphate = ADP + 1-phosphatidyl-1D-myo-inositol 4,5-
 CC bisphosphate.
 CC -!- SUBUNIT: Homodimer (By similarity).
 CC -!- TISSUE SPECIFICITY: EXPRESSED UBIQUITOUSLY, WITH HIGH LEVELS IN
 CC THE BRAIN.
 CC -!- SIMILARITY: BELONGS TO THE PTDINS(4)P-5-KINASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U14957; AAA64835.2; -
 DR EMBL; BC018034; AAH18034.1; -
 DR Genew; HGNC:8997; PIP5K2A.
 DR MIN; 603140; -
 DR InterPro; IPR002498; PIP5K.
 DR Pfam; PF01504; PIP5K; 1.
 KW Transferase; Kinase.
 SQ SEQUENCE 406 AA; 46224 MW; 5BAF0A27CC9EF376 CRC64;
 Query Match 55.4%; Score 36; DB 1; Length 406;
 Best Local Similarity 75.0%; Pred. No. 72;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 4 DSSGTHPV 11
 :|||
 DB 303 ESDGTHPV 310
 RESULT 25
 ID PI53 HUMAN STANDARD; PRT; 406 AA.
 AC P53807;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Phosphatidylinositol-4-phosphate 5-kinase type III (EC 2.7.1.149) (1-
 DE phosphatidylinositol-4-phosphate kinase) (PIP5KIII) (PtdIns(4)P-5-
 DE kinase C isoform) (Diphosphoinositide kinase).
 DE

GN PIP5K3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leukocyte;
RX MEDLINE=95366942; PubMed=7639683;
RA Divecha N., Truong O., Hsuan J.J., Hinchliffe K.A., Irvine R.F.;
RT "The cloning and sequence of the C isoform of Ptdins4P 5-kinase.";
RL Biochem. J. 309:715-719(1995).
CC -!- FUNCTION: Catalyzes the phosphorylation of phosphatidylinositol-4-
CC phosphate on the fifth hydroxyl of the myo-inositol ring, to form
CC phosphatidylinositol-4,5-bisphosphate.
CC -!- CATALYTIC ACTIVITY: ATP + 1-phosphatidyl-1D-myo-inositol 5-
CC phosphate = ADP + 1-phosphatidyl-1D-myo-inositol 4,5-bisphosphate.
CC -!- TISSUE SPECIFICITY: PARTICULARLY ABUNDANT IN PLATELETS AND IN
CC BRAIN. PRESENT IN MOST TISSUES, EXCEPT NOTABLY SKELETAL MUSCLE AND
CC SMALL INTESTINE.
CC -!- SIMILARITY: BELONGS TO THE PTDINS(4)P-5-KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@eb-sib.ch).
CC -----
DR EMBL; S78798; AB335041.1; -.
DR PIR; S57217; S57217.
DR GO; GO:0016308; F:1-phosphatidylinositol-4-phosphate 5-kinase. . .; NAS.
DR GO; GO:0016310; P:phosphorylation; NAS.
DR InterPro; IPR002498; PIP5K.
DR Pfam; PF01504; PIP5K; 1.
DR SMART; SM00330; FIPKC; 1.
KW Transferase; Kinase.
SQ SEQUENCE 406 AA; 46078 MW; E8DDFAF61A17534B CRC64;

Query Match 55.4%; Score 36; DB 1; Length 406;
Best Local Similarity 75.0%; Pred. No. 72;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 DSSGTHPV 11
Db 303 ESDGTHPV 310

Search completed: September 27, 2004, 12:13:55
Job time : 28 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 27, 2004, 12:07:49 ; Search time 112 Seconds
(without alignments)
30.988 Million cell updates/sec

Title: US-10-088-639A-2_COPY_88_98
Perfect score: 65
Sequence: 1 NSWDSGTHPV 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database :

SPTREMBL_25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_orphan:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	43	66.2	234	4	Q8N355 homo sapien
2	43	66.2	508	4	Q9H797 homo sapien
3	43	66.2	508	4	Q96F86 homo sapien
4	43	66.2	508	11	Q9K2D3
5	42	64.6	130	2	Q56910
6	42	64.6	669	13	Q98TU7
7	41	63.1	98	16	Q97FD1
8	41	63.1	339	5	Q23547
9	41	63.1	456	16	Q98QW0
10	41	63.1	1461	10	Q94H87
11	40.5	62.3	111	11	Q9DBS6
12	40	61.5	261	5	Q9WSV6
13	40	61.5	374	2	Q9S420
14	40	61.5	469	16	Q9G7R9
15	40	61.5	824	2	Q45464
16	40	61.5	914	10	P93659

17	40	61.5	988	10	P93065
18	39	60.0	60	5	Q9BIA7
19	39	60.0	230	4	Q8N4H3
20	39	60.0	231	16	Q82023
21	39	60.0	250	4	Q8N316
22	39	60.0	254	16	Q8N3N0
23	39	60.0	270	10	Q40987
24	39	60.0	280	4	Q8IW63
25	39	60.0	285	10	Q8S860
26	39	60.0	301	10	Q7XKX8
27	39	60.0	305	4	Q8N9A5
28	39	60.0	305	4	Q8N386
29	39	60.0	347	4	Q96H00
30	39	60.0	357	5	Q9BIA9
31	39	60.0	394	16	Q9F2L1
32	39	60.0	402	16	P72675
33	39	60.0	422	16	Q8CX87
34	39	60.0	482	3	Q7Z962
35	39	60.0	593	13	Q9PUC2
36	39	60.0	596	16	Q8RB15
37	39	60.0	621	10	Q949J9
38	39	60.0	627	10	Q8LJ95
39	39	60.0	652	10	Q8SX70
40	39	60.0	937	16	Q82PD9
41	39	60.0	995	5	P90897
42	39	60.0	1336	10	Q81074
43	38	58.5	53	4	Q96GB3
44	38	58.5	107	4	Q9NSD6
45	38	58.5	163	2	Q93IH5
46	38	58.5	295	10	Q9ZW81
47	38	58.5	398	16	Q9A1Z8
48	38	58.5	398	16	Q8P318
49	38	58.5	471	2	Q93T91
50	38	58.5	482	2	Q9FAC7
51	38	58.5	490	2	Q33853
52	38	58.5	541	16	Q87EE0
53	38	58.5	567	16	Q97ED4
54	38	58.5	620	15	Q99212
55	38	58.5	652	4	Q86TG4
56	38	58.5	943	11	Q8BK12
57	38	58.5	958	5	Q95PM8
58	38	58.5	978	5	Q7YVL7
59	38	58.5	1099	10	Q9XFK8
60	38	58.5	1139	12	Q65714
61	38	58.5	1369	10	Q23027
62	38	58.5	1390	4	Q86V60
63	38	58.5	1755	17	Q26812
64	38	58.5	1884	4	Q86VW8
65	38	58.5	1887	11	Q9QW67
66	38	58.5	1898	11	Q9EQ17
67	38	58.5	1898	11	Q64604
68	37	56.9	22	3	Q9URC3
69	37	56.9	22	3	Q9URC2
70	37	56.9	81	4	Q7Z2B8
71	37	56.9	107	2	Q7X1H7
72	37	56.9	125	16	Q9KFP4
73	37	56.9	170	16	Q8E966
74	37	56.9	236	2	Q9L920
75	37	56.9	257	5	Q9VNY6
76	37	56.9	271	16	Q9CDC5
77	37	56.9	274	11	Q8C089
78	37	56.9	292	6	Q9XT99
79	37	56.9	295	4	Q95603
80	37	56.9	295	4	Q86U58
81	37	56.9	295	6	Q95KM3
82	37	56.9	295	6	Q29476
83	37	56.9	307	10	Q8H7Y9
84	37	56.9	329	4	Q8N8F1
85	37	56.9	330	4	Q81YF4
86	37	56.9	330	11	Q8C197
87	37	56.9	372	2	Q45559
88	37	56.9	372	2	Q68884
89	37	56.9	375	2	Q59419

P93065	brassica na
Q9BIA7	caenorhabdi
Q8N4H3	homo sapien
Q82023	tropheryma
Q8N316	homo sapien
Q8N3N0	tropheryma
Q40987	pisum sativ
Q8IW63	homo sapien
Q8S860	oryza sativ
Q7XKX8	oryza sativ
Q8N9A5	homo sapien
Q8N386	homo sapien
Q96H00	homo sapien
Q9BIA9	caenorhabdi
Q9F2L1	streptomyce
P72675	synecocyst
Q8CX87	leptospira
Q7Z962	coccidioid
Q9PUC2	brachydanio
Q8RB15	thermoanaer
Q949J9	lycopersico
Q8LJ95	oryza sativ
Q8SX70	arabidopsis
Q82PD9	streptomyce
P90897	caenorhabdi
Q81074	arabidopsis
Q96GB3	homo sapien
Q9NSD6	homo sapien
Q93IH5	wolinella s
Q9ZW81	arabidopsis
Q9A1Z8	streptococc
Q8P318	streptococc
Q93T91	azospirillu
Q9FAC7	rhodococcus
Q33853	rhodococcus
Q87FE0	vibrio para
Q97ED4	clostridium
Q99212	mus cervico
Q86TG4	homo sapien
Q8BK12	mus musculu
Q95PM8	paracentrot
Q7YVL7	trypanosoma
Q9XFK8	zea mays (m
Q65714	barley stri
Q23027	arabidopsis
Q86V60	homo sapien
Q26812	methanobact
Q86VW8	homo sapien
Q9QW67	rattus sp.
Q9EQ17	mus musculu
Q64604	r protein-t
Q9URC3	phanerocha
Q9URC2	phanerocha
Q7Z2B8	homo sapien
Q7X1H7	leptospiril
Q9KFP4	bacillus ha
Q8E966	shewanella
Q9L920	streptococc
Q9VNY6	drosophila
Q9CDC5	mycobacteri
Q8C089	mus musculu
Q9XT99	oryctolagus
Q95603	homo sapien
Q86U58	homo sapien
Q95KM3	oryctolagus
Q29476	canis famli
Q8H7Y9	oryza sativ
Q8N8F1	homo sapien
Q81YF4	homo sapien
Q8C197	mus musculu
Q45559	bacillus su
Q68884	bacillus su
Q59419	erwinia chr

90 37 56.9 394 16 006164
 91 37 56.9 394 16 07TYF7
 92 37 56.9 442 13 073803
 93 37 56.9 511 8 Q85ZS6
 94 37 56.9 515 16 087IC6
 95 37 56.9 567 16 081P81
 96 37 56.9 567 16 081CL9
 97 37 56.9 626 16 09KY10
 98 37 56.9 661 16 08D7M5
 99 37 56.9 689 10 07XIH7
 100 37 56.9 760 10 08SOA4
 101 37 56.9 771 16 087IC9
 102 37 56.9 873 16 08PD52
 103 37 56.9 1058 16 08LP18
 104 37 56.9 1606 4 Q81WC7
 105 37 56.9 1622 10 048908
 106 37 56.9 1623 10 064590
 107 37 56.9 1623 10 022449
 108 37 56.9 1623 10 048907
 109 37 56.9 1732 10 023576
 110 37 56.9 2013 16 092EK2
 111 37 56.9 2013 16 08Y9T8
 112 36.5 56.2 336 12 07T928
 113 36 55.4 49 Q8TX33
 114 36 55.4 108 16 087CU6
 115 36 55.4 150 17 09YE74
 116 36 55.4 186 10 09FFV2
 117 36 55.4 192 9 Q9MC44
 118 36 55.4 201 16 088KL8
 119 36 55.4 226 15 086833
 120 36 55.4 247 15 086834
 121 36 55.4 251 11 08C486
 122 36 55.4 256 16 07VXD0
 123 36 55.4 259 16 07WJ61
 124 36 55.4 259 16 07WA29
 125 36 55.4 272 16 082DY2
 126 36 55.4 289 16 098IT3
 127 36 55.4 328 4 Q95061
 128 36 55.4 336 2 Q44128
 129 36 55.4 379 2 Q8RTY0
 130 36 55.4 380 16 082SP1
 131 36 55.4 394 16 08YXV4
 132 36 55.4 395 11 Q9EPK5
 133 36 55.4 395 11 Q99KI4
 134 36 55.4 398 4 Q8N3P2
 135 36 55.4 399 8 Q9G8U1
 136 36 55.4 400 4 Q9GZV5
 137 36 55.4 404 16 082L46
 138 36 55.4 406 6 Q13010
 139 36 55.4 424 5 Q76744
 140 36 55.4 426 10 Q23313
 141 36 55.4 450 2 Q52869
 142 36 55.4 451 10 Q9LMH7
 143 36 55.4 456 16 082JP7
 144 36 55.4 469 16 Q9ASU9
 145 36 55.4 475 4 Q96S29
 146 36 55.4 494 10 Q852N8
 147 36 55.4 526 5 Q8MLC7
 148 36 55.4 550 16 Q8YR96
 149 36 55.4 620 16 Q98Q29
 150 36 55.4 636 4 Q9NW05

ALIGNMENTS

RESULT 1
 Q8N355 PRELIMINARY; PRT; 234 AA.
 ID Q8N355;
 AC Q8N355;
 DT 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)

DE Hypothetical protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Strausberg R.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC028090; AAH28090.1; -.
 DR PIR; S12441; S12441.
 DR InterPro; IPR003599; IG.
 DR InterPro; IPR007110; IG-like.
 DR InterPro; IPR003597; IG cl.
 DR InterPro; IPR003006; IG_MHC.
 DR InterPro; IPR003596; IG_v.
 DR Pfam; PF00047; IG; 2.
 DR SMART; SM00409; IG; 2.
 DR SMART; SM00407; IGL; 1.
 DR SMART; SM00406; IGV; 1.
 DR PROSITE; PS00835; IG LIKE; 2.
 DR PROSITE; PS00290; IG_MHC; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 234 AA; 24792 MW; CC848CAEBA4A9D63 CRC64;

Query Match 66.2%; Score 43; DB 4; Length 234;
 Best Local Similarity 77.8%; Pred. No. 22;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 WSSSGTHPV 11
 Db 109 WSSSDHPV 117

RESULT 2

QSH797 PRELIMINARY; PRT; 508 AA.
 ID Q9H797
 AC Q9H797;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Hypothetical protein FLJ21128.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,
 RA Suzuki Y., Ohtsuka M., Nishi T., Shibahara T., Tanaka T.,
 RA Nakamura Y., Isegai T., Sugano S.;
 RT "NEDO human cDNA sequencing project."
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK024781; BAB15001.1; -.
 DR InterPro; IPR004443; YJef_Nterm.
 DR Pfam; PF03853; YJef_N; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 508 AA; 56105 MW; AAFABFD0222E23DB CRC64;

Query Match 66.2%; Score 43; DB 4; Length 508;
 Best Local Similarity 70.0%; Pred. No. 51;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 NSWDSSGTHP 10
 Db 160 NSWSSSRHP 169

RESULT 3

Q96F86 PRELIMINARY; PRT; 508 AA.
 ID Q96F86
 AC Q96F86;

DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN PP844.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A.
RP TISSUE=Placenta, and Eye;
RC Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Gu J.R., Wan D.F., Zhao X.T., Zhou X.M., Jiang H.Q., Zhang P.P.,
RA Qin W.X., Huang Y., Qiu X.K., Qian L.F., He L.P., Li H.N., Yu Y.,
RA Yu J., Han L.H.;
RT "Novel human cDNA clones with function of inhibiting cancer cell
RT growth."
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC011534; AAH11534.1; -;
DR EMBL; BC021271; AAH21271.1; -;
DR EMBL; AF193058; AAG22486.1; -;
DR InterPro; IPR004443; Yjef_Nterm.
DR Pfam; PF03853; Yjef_N; 1.
KW Hypothetical protein.
SQ SEQUENCE 508 AA; 56077 MW; 95346F484FCF3EA CRC64;
Query Match 66.2%; Score 43; DB 4; Length 508;
Best Local Similarity 70.0%; Pred. No. 51;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 NSWSSSGTHP 10
DB 160 NSWSSSSRRHP 169
RESULT 4
Q8K2D3 PRELIMINARY; PRT; 508 AA.
AC Q8K2D3;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Similar to hypothetical protein FLJ21128.
GN CLK3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC031725; AAH31725.1; -;
DR EMBL; BC033484; AAH33484.1; -;
DR MGD; MGI:1098670; Clk3.
DR InterPro; IPR004443; Yjef_Nterm.
DR Pfam; PF03853; Yjef_N; 1.
KW Hypothetical protein.
SQ SEQUENCE 508 AA; 55957 MW; DD1F449BEF1341C8 CRC64;
Query Match 66.2%; Score 43; DB 11; Length 508;
Best Local Similarity 70.0%; Pred. No. 51;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 NSWSSSGTHP 10
DB 160 NSWSSSSRRHP 169

RESULT 5
Q56910 PRELIMINARY; PRT; 130 AA.
AC Q56910;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE YopE chaperone Syce.
GN SYCE.
OS Yersinia enterocolitica.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=630;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SW22703;
RA Wattiau P., Cornelis G.R.;
RT "Syce, a chaperone-like protein of Yersinia enterocolitica involved in
RT the secretion of YopE."
RL Mol. Microbiol. 8:123-131(1993).
DR EMBL; AF102990; AAD16849.1; -;
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
DR InterPro; IPR005416; Syce Chap.
DR PRINTS; PR01596; SYCECHAPRONE.
KW Plasmid.
SQ SEQUENCE 130 AA; 14649 MW; 490B89BA3293CASA CRC64;
Query Match 64.6%; Score 42; DB 2; Length 130;
Best Local Similarity 70.0%; Pred. No. 17;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2 SWDSGTHPV 11
DB 79 SWDEVGHPV 88
RESULT 6
Q98TU7 PRELIMINARY; PRT; 669 AA.
AC Q98TU7;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Serine kinase protein.
OS Agelaius phoeniceus (Red-winged blackbird).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Passeriformes; Icteridae; Agelaius.
OX NCBI_TaxID=39638;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=21365706; PubMed=11472064;
RA Gasper J.S., Shiina T., Inoko H., Edwards S.V.;
RT "Songbird genomics: analysis of 45 kb upstream of a polymorphic Mhc
RT class II gene in red-winged blackbirds (Agelaius phoeniceus).";
RL Genomics 75:26-34(2001).
DR EMBL; AF328738; AAK08500.1; -;
DR HSSP; P24941; 1BUH.
DR GO; GO:0005524; P:ATP binding; IEA.
DR GO; GO:0004674; P:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; P:protein-tyrosine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SMO0220; S_TK; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.

```

KW ATP-binding; Kinase; Transferase.
SQ SEQUENCE 669 AA; 73710 MW; 5987154DCB1C441A CRC64;

  Query Match      64.6%; Score 42; DB 13; Length 669;
  Best Local Similarity 66.7%; Pred. No. 1e+02; 3; Indels 0; Gaps 0;
  Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 WDSGTHPV 11
Db 170 WDEGAEHV 178

RESULT 7
Q97FD1 PRELIMINARY; PRT; 98 AA.
AC Q97FD1;
DT 01-OCT-2001 (TREMBLrel. 18, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Predicted HD superfamily hydrolase.
GN CAC2809.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1488;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=1146286; M.V., Makarova K.S., Zeng Q.,
RA Neelling J., Breton G., Omelchenko M.V., Hitti J., Wolf Y.I.,
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA Tatusov R.N., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA Bennett G.N., Koonin E.V., Smith D.R.;
RA "Genome sequence and comparative analysis of the solvent-producing
  bacterium Clostridium acetobutylicum."
RT J. Bacteriol. 183:4823-4838(2001).
RL J. Bacteriol. 183:4823-4838(2001).
DR EMBL; AE007778; AAK80753.1; -.
DR PIR; F97245; F97245.
DR GO; GO:0016787; F:hydrolase activity; IEA.
KW Hydrolase; Complete proteome.
SQ SEQUENCE 98 AA; 11692 MW; 929C71F938AC930A CRC64;

  Query Match      63.1%; Score 41; DB 16; Length 98;
  Best Local Similarity 77.8%; Pred. No. 19;
  Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 SWDSGTHP 10
Db 40 SWDQSGTTP 48

RESULT 8
Q23547 PRELIMINARY; PRT; 339 AA.
ID Q23547;
AC Q23547;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE ZK596.3 protein.
GN ZK596.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA McMurray A.A.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RA "Genome sequence of the nematode C.elegans: A platform for

```

```

RT investigating biology."
RL Science 282:2012-2018(1998).
DR EMBL; Z69386; CAA93432.1; -.
DR PIR; T27931; T27931.
DR WormPep; ZK596.3; CE06633.
SQ SEQUENCE 339 AA; 39145 MW; 21CA0504F4CE1844 CRC64;

  Query Match      63.1%; Score 41; DB 5; Length 339;
  Best Local Similarity 70.0%; Pred. No. 72;
  Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 SWDSGTHPV 11
Db 48 SWNSNGTHYV 57

RESULT 9
Q98QW0 PRELIMINARY; PRT; 456 AA.
ID Q98QW0;
AC Q98QW0;
DT 01-OCT-2001 (TREMBLrel. 18, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Hypothetical protein MYPU_2500.
GN MYPU_2500.
OS Mycoplasma pulmonis.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2107;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UAB CTIP;
RX MEDLINE=21267165; PubMed=1153084;
RA Chambaud I., Heilig R., Ferris S., Barbe V., Samson D., Galisson F.,
RA Moszer I., Dybvig K., Wroblewski H., Viari A., Rocha E.P.C.,
RA Blanchard A.;
RA "The complete genome sequence of the murine respiratory pathogen
  Mycoplasma pulmonis."
RT Nucleic Acids Res. 29:2145-2153(2001).
RL EMBL; AL445563; CAC13423.1; -.
DR PIR; B90543; B90543.
DR MYPUlist; MYPU 2500; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 456 AA; 53149 MW; 889D356B34946FB5 CRC64;

  Query Match      63.1%; Score 41; DB 16; Length 456;
  Best Local Similarity 70.0%; Pred. No. 99;
  Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NSWDSSGTHP 10
Db 19 NSWDISKTP 28

RESULT 10
Q94H87 PRELIMINARY; PRT; 1461 AA.
ID Q94H87;
AC Q94H87;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Putative disease resistance protein.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Buell C.R., Yuan Q., Ouyang S., Moffat K.S., Hill J.N., Gansberger K.,
RA Brenner M., Burgess S., Hance M., Shvartsbeyn M., Tsitrin T.,
RA Riggs F., Hsiao J., Ziemann V., Blunt S., Pai G., VanAken S.E.,
RA Utterback T.R., Feldblyum T.V., Quackenbush J., Salzberg S.L.,

```

```

RA White O., Fraser C.M.;
RT "Oryza sativa chromosome 3 BAC OSJNBa0090P23 genomic sequence.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC084380; AAK52137.1; -
DR Gramene; Q94H87; -
DR InterPro; IPR001611; LRR-
DR InterPro; IPR007090; LRR plant.
DR InterPro; IPR002885; PPR.
DR InterPro; IPR008941; TPR-like.
DR Pfam; PF00560; LRR: 14.
DR Pfam; PF01535; PPR: 5.
DR PRINTS; PR00019; LEURICHRPT.
DR TIGRFAms; TIGR00756; PPR: 5.
SQ SEQUENCE 1461 AA; 158782 MW; 9A89A942B561D9F2 CRC64;

Query Match 63.1%; Score 41; DB 10; Length 1461;
Best Local Similarity 70.0%; Pred. No. 3.5e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 NSWDSSGTHP 10
Db 49 NSWPESGTSP 58

RESULT 11
ID Q9DB56 PRELIMINARY; PRT; 111 AA.
AC Q9DB56;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE 1500010020Rik protein.
GN 1500010020Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN JALALI M., Kallush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Giess C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staib F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Guscinich S., Hall D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690 (2001).
DR EMBL; AK005197; BAB23877.1; -
DR MGD; MGI:1914953; 1500010020Rik.
SQ SEQUENCE 111 AA; 12221 MW; 1BC7884BFP056C43B CRC64;

Query Match 62.3%; Score 40.5; DB 11; Length 111;
Best Local Similarity 66.7%; Pred. No. 26;
Matches 8; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

Qy 1 NSWDSSG-THPV 11
Db 59 NSWRRTGTHPV 70

```

```

RESULT 12
Q9WSV6
ID Q9WSV6 PRELIMINARY; PRT; 261 AA.
AC Q9WSV6; Q9WSV5;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE CG12414 protein.
GN NACE-ALPHA-80B OR NACRALPHA-80B OR CG12414 OR CG17552.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celnikier S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.M., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., B.D.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
RA Foster K.J., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kallush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassearman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Celnikier S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busan D.A.,
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA Ferraz C., Frise E., Galle R.F., Garg N.S., George R.A.,
RA Gonzalez S., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA Pacle J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT "Sequencing of Drosophila melanogaster genome.";

```

```

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA Tupy J.L., Bergman C., Bernan B., Carlson J.W., Celniker S.E.,
RA Clump A.M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA Ashburner M., Gelbart W.M., Rubin G.M., Muncall C.J., Lewis S.E.;
RT "Annotation of Drosophila melanogaster genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE002665; AAF45409.2; -.
DR FlyBase; FBgn0037212; nacr-alpha-80B.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005230; F:extracellular ligand-gated ion channel acti. . .; IEA.
DR GO; GO:0030594; F:neurotransmitter receptor activity; IEA.
DR GO; GO:0006811; P:ion transport; IEA.
DR InterPro; IPR006202; Neur_chan_LBD.
DR InterPro; IPR006029; Neur_chan_memb.
DR Pfam; PF02931; Neur_chan_LBD; 1.
DR Pfam; PF02932; Neur_chan_memb; 1.
SQ SEQUENCE 261 AA; 30019 MW; 4DGB67F85F385FC CRC64;

Query Match 61.5%; Score 40; DB 5; Length 261;
Best Local Similarity 60.0%; Pred. No. 80;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 SWDSSGTHPV 11
DB 252 SWPSAATHPL 261

RESULT 13
Q59420
ID Q59420 PRELIMINARY; PRT; 374 AA.
AC Q59420;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-AUG-1999 (TrEMBLrel. 11, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Pectate lyase precursor (EC 4.2.2.2).
GN PELC.
OS Erwinia chrysanthemi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pectobacterium.
OX NCBI_TaxID=556;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3937;
RA Hugouvieux-Cotte-Pattat N.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 1-54 FROM N.A.
RC STRAIN=3937;
RX MEDLINE=93023864; PubMed=1406275;
RA Hugouvieux-Cotte-Pattat N., Robert-Baudouy J.;
RT "Analysis of the regulation of the pelBC genes in Erwinia chrysanthemi
RT 3937.";
RL Mol. Microbiol. 6:2363-2376(1992).
CC -!- CATALYTIC ACTIVITY: ELIMINATIVE CLEAVAGE OF PECTATE TO GIVE
CC OLIGOSACCHARIDES WITH 4-DEOXY-ALPHA-D-GLUC-4-ENURONOSYL GROUPS AT
CC THEIR NON-REDUCING ENDS.
DR EMBL; AJ132325; CAA10642.1; -.
DR EMBL; X67475; CAA47822.1; -.
DR PIR; S25263; S25263.

DR HSPF; P11073; IAIR.
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0030570; F:pectate lyase activity; IEA.
DR InterPro; IPR002022; Amb_allergen.
DR Pfam; PF00544; Pec_lyase; 1.
DR SMART; SM00656; Amb_all; 1.
DR Lyase; Signal.
FT CHAIN 1 22 POTENTIAL.
FT CHAIN 23 374 PECTATE LYASE.
SQ SEQUENCE 374 AA; 39891 MW; 0086E71012E13FD4 CRC64;

Query Match 61.5%; Score 40; DB 2; Length 374;
Best Local Similarity 60.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 NSWSSSGTHP 10
DB 328 DSWTSTGTYP 337

RESULT 14
Q8G7R9
ID Q8G7R9 PRELIMINARY; PRT; 469 AA.
AC Q8G7R9;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Possible alpha-galactosidase.
GN BL0177.
OS Bifidobacterium longum.
OC Bacteria; Actinobacteria; Actinobacteridae; Bifidobacteriales;
OC Bifidobacteriaceae; Bifidobacterium.
OX NCBI_TaxID=216816;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCC.2705;
RX MEDLINE=22294977; PubMed=12381787;
RA Schell M.A., Karmirantzou M., Snel B., Vilanova D., Berger B.,
RA Pessi G., Zwahlen M.-C., Desiere F., Bork P., Delley M.,
RA Pridmore R.D., Arigoni F.;
RT "The genome sequence of Bifidobacterium longum reflects its adaptation
RT to the human gastrointestinal tract.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14422-14427(2002).
DR EMBL; AE014635; AAN24032.1; -.
DR GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . .; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR002241; Glyco_hydro_27.
DR InterPro; IPR000111; Glyco_hydro_GHD.
DR PRINTS; PR00740; GLYDRLASE27.
DR ProDom; PD002572; Glyco_hydro_GHD; 1.
KW Complete proteome.
SQ SEQUENCE 469 AA; 52574 MW; 6A9A36769ABAAAA3 CRC64;

Query Match 61.5%; Score 40; DB 16; Length 469;
Best Local Similarity 87.5%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NSWSSSGT 8
DB 19 NSWDSYGT 26

RESULT 15
Q45464
ID Q45464 PRELIMINARY; PRT; 824 AA.
AC Q45464; Q45465;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE SPRB.
OS Bacillus sp.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1409;

```

```

RN RP SEQUENCE FROM N.A.
RC STRAIN=LG12;
RX MEDLINE=96086035; PubMed=8534117;
RA Schmidt B.P., Woodhouse L., Adams R.M., Ward T., Mainzer S.E.,
RA Lad P.J.;
RT "Alkalophilic Bacillus sp. strain LG12 has a series of serine protease
genes.";
RL Appl. Environ. Microbiol. 61:4490-4493(1995).
RC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
DR EMBL; U39230; AAC43578.1; -.
DR EMBL; U39230; AAC43579.1; -.
DR HSSP; Q45670; IDBI.
DR GO; GO:0008223; F:peptidase activity; IEA.
DR GO; GO:0004289; F:subtilase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000209; Peptidase_S8.
DR InterPro; IPR001119; SLH.
DR Pfam; PF00082; Peptidase_S8; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
DR Hydrolase; Protease; Serine protease.
KW SEQUENCE 824 AA; 88338 MW; 0D655CB86AFB5CFF CRC64;

Query Match 61.5%; Score 40; DB 2; Length 824;
Best Local Similarity 53.3%; Pred. No. 2.8e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 4; Gaps 1;

Qy 1 NSWDSSG-----THPV 11
Db 273 NSWGSGGDFDPNHP 287

RESULT 16
P93659
ID P93659 PRELIMINARY; PRT; 914 AA.
AC P93659;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Jasmonate inducible protein.
OS Brassica napus (Rape).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids II; Brassicales; Brassicaceae;
OX NCBI_TaxID=3708;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Global;
RA Geisli N., Brandt A.;
RT "Two jasmonate inducible proteins from Brassica napus seedlings
RT homologous to myrosinase binding proteins and jacalin.";
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; Y11482; CAA72270.1; -.
DR PIR; T08081; T08081.
DR HSSP; P18670; LJAC.
DR InterPro; IPR001229; Jacalin_lectin.
DR Pfam; PF01419; Jacalin; 6.
SQ SEQUENCE 914 AA; 96911 MW; 9126F5BB2482155F CRC64;

Query Match 61.5%; Score 40; DB 10; Length 914;
Best Local Similarity 66.7%; Pred. No. 3.1e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 NSWDSSGTH 9
Db 683 NQWDDSGDH 691

RESULT 17
P93065

```

```

ID P93065 PRELIMINARY; PRT; 988 AA.
AC P93065;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Myrosinase binding protein.
OS Brassica napus (Rape).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids II; Brassicales; Brassicaceae;
OX NCBI_TaxID=3708;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leaf;
RX MEDLINE=98121188; PubMed=9461290;
RA Taipaleenau J., Eriksson S.J., Raek L.;
RT "The myrosinase binding protein from Brassica napus seeds possesses
RT lectin activity and has a highly similar vegetatively expressed wound
RT inducible counterpart.";
RL Eur. J. Biochem. 250:680-688(1997).
DR EMBL; Y09437; CAA70587.1; -.
DR PIR; T08102; T08102.
DR HSSP; P18674; LJOT.
DR InterPro; IPR001229; Jacalin_lectin.
DR Pfam; PF01419; Jacalin; 6.
SQ SEQUENCE 988 AA; 104331 MW; 6C1DB595F56F246C CRC64;

Query Match 61.5%; Score 40; DB 10; Length 988;
Best Local Similarity 66.7%; Pred. No. 3.4e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 NSWDSSGTH 9
Db 699 NQWDDSGDH 707

RESULT 18
P93065
ID P93065 PRELIMINARY; PRT; 60 AA.
AC P93065;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN C17C3.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Du Z.;
RT "The sequence of C. elegans cosmid C17C3.";
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U41279; AAK31432.1; -.
DR WormPep; C17C3.1c; CE27072.
KW Hypothetical protein.
SQ SEQUENCE 60 AA; 7054 MW; CA8069E3967A80F0 CRC64;

```

Query Match 60.0%; Score 39; DB 5; Length 60;
 Best Local Similarity 54.5%; Pred. No. 24;
 Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 NSWSSGTHPV 11
 Db 20 NSWEVPSNHPV 30

RESULT 19

Q8N4H3 Q8N4H3 PRELIMINARY; PRT; 230 AA.
 AC Q8N4H3;
 DT 01-OCT-2002 (TReMBLrel. 22, Created)
 DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
 DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
 DE Similar to RIKEN CDNA 290006N09 gene.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Strausberg R.;
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC034275; AAH34275.1; -.
 DR InterPro; IPR007803; Asp_Arg_Hydrox.
 DR Pfam; PF05118; Asp_Arg_Hydrox; 1.
 SQ SEQUENCE 230 AA; 24622 MW; 30FC26C4B4F991F CRC64;

Query Match 60.0%; Score 39; DB 4; Length 230;
 Best Local Similarity 77.8%; Pred. No. 1e+02;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 SWDSSGTHP 10
 Db 68 SWDFSCTTP 76

RESULT 20

Q82023 Q82023 PRELIMINARY; PRT; 231 AA.
 AC Q82023;
 DT 01-JUN-2003 (TReMBLrel. 24, Created)
 DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
 DE Conserved hypothetical protein (Possible ATP-binding).
 GN TW138.
 OS Tropheryma whipplei (strain TW08/27) (Whipple's bacillus).
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Micrococcineae; Cellulomonadaceae; Tropheryma.
 OX NCBI_TaxID=218496;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22495039; PubMed=13606174;
 RA Bentley S.D., Maitwald M., Murphy L.D., Pallen M.J., Yeats C.A.,
 RA Dover L.G., Norbertzak H.T., Besra G.S., Quail M.A., Harris D.E.,
 RA von Herbay B.G., Goble A., Rutter S., Squares R., Squares S.,
 RA Barrell B.G., Parkhill J., Rellman D.A.;
 RT "Sequencing and analysis of the genome of the Whipple's disease
 bacterium Tropheryma whipplei.",
 RL Lancet 361:637-644(2003).
 DR EMBL; BX251410; CAD66818.1; -.
 DR GO; GO:0005524; F:ATP binding; IEA.
 KW Hypothetical protein; ATP-binding; Complete proteome.
 SQ SEQUENCE 231 AA; 26160 MW; FDCA42527C8664A2 CRC64;

Query Match 60.0%; Score 39; DB 16; Length 231;
 Best Local Similarity 70.0%; Pred. No. 1e+02;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 SWDSSGTHPV 11
 Db 100 SWRSRGFHPV 109

RESULT 21

Q8N316 Q8N316 PRELIMINARY; PRT; 250 AA.
 AC Q8N316;
 DT 01-OCT-2002 (TReMBLrel. 22, Created)
 DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
 DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
 DE Hypothetical protein (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Colon, and Kidney;
 RA Strausberg R.;
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC029136; AAH29136.1; -.
 DR InterPro; IPR007803; Asp_Arg_Hydrox.
 DR Pfam; PF05118; Asp_Arg_Hydrox; 1.
 KW Hypothetical protein.
 FT NON TER 1
 SQ SEQUENCE 250 AA; 27072 MW; 78CAB08AA4EF1E44 CRC64;

Query Match 60.0%; Score 39; DB 4; Length 250;
 Best Local Similarity 77.8%; Pred. No. 1.1e+02;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 SWDSSGTHP 10
 Db 88 SWDFSCTTP 96

RESULT 22

Q83N30 Q83N30 PRELIMINARY; PRT; 254 AA.
 AC Q83N30;
 DT 01-JUN-2003 (TReMBLrel. 24, Created)
 DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
 DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
 DE ComF operon protein 3.
 GN COMFC OR TW129.
 OS Tropheryma whipplei (strain Twist) (Whipple's bacillus).
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Micrococcineae; Cellulomonadaceae; Tropheryma.
 OX NCBI_TaxID=203267;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Raoult D., Audic S., Robert C., Ogata H., Suhre K., Drancourt M.,
 RA Claverie J.-M.;
 RT "Tropheryma whipplei illustrates the diversity of gene loss patterns
 in small genome bacterial pathogens.",
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE016850; AAO44226.1; -.
 KW Complete proteome.
 SQ SEQUENCE 254 AA; 28732 MW; 32C29A6F8E00FD0 CRC64;

Query Match 60.0%; Score 39; DB 16; Length 254;
 Best Local Similarity 70.0%; Pred. No. 1.2e+02;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 SWDSSGTHPV 11
 Db 123 SWRSRGFHPV 122

RESULT 23

Q40987

```

ID Q40987 PRELIMINARY; PRT; 270 AA.
AC Q40987;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Nodule lectin.
GN PSNLECL.
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.
OX NCBI_TaxID=3888;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Wisconsin Perfection; TISSUE=Root nodules;
RX MEDLINE=9222420; PubMed=8685275;
RA Kardalis I.V., Sherrier D.J., Brewin N.J.;
RT "Identification of a new pea gene, PnLec1, encoding a lectin-like
RT glycoprotein isolated from the symbiosomes of root nodules.";
RL Plant Physiol. 111:49-60(1996).
DR EMBL; U31981; AAC49367.1; -.
DR PIR; T06528; T06528.
DR HSSP; P05046; 1SBD.
DR GO; GO:0005529; F:sugar binding; IEA.
DR GO; GO:0007157; P:heterophilic cell adhesion; IEA.
DR InterPro; IPR008985; CoNA like lec_gl.
DR InterPro; IPR000985; Lectin_legA.
DR InterPro; IPR001220; Lectin_legB.
DR Pfam; PF00138; lectin_legA; 1.
DR Pfam; PF00139; lectin_legB; 1.
DR ProDom; PD000671; Lectin_legA; 1.
DR ProDom; PD000711; Lectin_legB; 1.
KW Lectin.
SQ SEQUENCE 270 AA; 28958 MW; 648D9CDBF8E5E2E3 CRC64;

Query Match 60.0%; Score 39; DB 10; Length 270;
Best Local Similarity 66.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 NSWDSSGTHP 9
Db 162 NSWDPEGRH 170

RESULT 24
ID Q81W63 PRELIMINARY; PRT; 280 AA.
AC Q81W63;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC040933; AAH40933.1; -.
DR InterPro; IPR007803; Asp_Arg_Hydrox.
DR Pfam; PF05118; Asp_Arg_Hydrox; 1.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 280 AA; 29753 MW; CAA2A3804ACCI960 CRC64;

Query Match 60.0%; Score 39; DB 4; Length 280;
Best Local Similarity 77.8%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 SWDSSGTHP 10

```

```

Db 118 SWDFSGTTP 126

RESULT 25
Q8S860
ID Q8S860 PRELIMINARY; PRT; 285 AA.
AC Q8S860;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN OSJNBA0061K21.17.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Spiegel L.A., Nascimento L.U., de la Bastide M., Vil D.M.,
RA Preston R.R., Huang E.N., See L.H., Shah R.S., O'Shaughnessy A.,
RA Rodriguez M.A., Shekher M., Kirchoff K.A., Baker J.P., Schutz K.,
RA Dedhia N.N., McCombie W.R.;
RT "Genomic Sequence For Oryza sativa, Niponbare Strain, Clone
RT OSJNBA0061K21 From Chromosome 10, Complete Sequence.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA McCombie W.R.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA McCombie W.R.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Palmer L.E., Yu M., Spiegel L.A., Nascimento L.U., de la Bastide M.,
RA Vil D.M., Preston R.R., Huang E.N., See L.H., Shah R.S.,
RA O'Shaughnessy A., Rodriguez M.A., Shekher M., Kirchoff K.A.,
RA Baker J.P., Schutz K., Dedhia N.N., McCombie W.R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC016780; AAC08786.1; -.
DR Gramene; Q8S860; -.
DR InterPro; IPR006502; DUF_plant_1615.
DR Pfam; PF04720; DUF506; 1.
DR TIGRFAMs; TIGR01615; A_thal_3542; 1.
KW Hypothetical protein.
SQ SEQUENCE 285 AA; 30493 MW; C863AB612D5151B3 CRC64;

Query Match 60.0%; Score 39; DB 10; Length 285;
Best Local Similarity 75.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 WDSGTHP 10
Db 151 WDXSPTH 158

Search completed: September 27, 2004, 12:15:56
Job time : 119 secs

```

This Page Blank (uspio)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2005, 17:57:45 ; Search time 71.1562 Seconds
(without alignments)
79.162 Million cell updates/sec

Title: US-10-088-639A-2_COPY_88_98
Perfect score: 65
Sequence: 1 NSWDSSGTHPV 11

Scoring table: BLOSUM62
Gap 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03: *
1: uniprot_spot: *
2: uniprot_trembl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	43	66.2	234	2 Q8N355	Q8N355 homo sapien
2	43	66.2	508	2 Q96F86	Q96F86 homo sapien
3	43	66.2	508	2 Q9H797	Q9H797 homo sapien
4	43	66.2	508	2 Q8K2D3	Q8K2D3 mus musculus
5	42	64.6	130	1 YERA_YEREN	P31490 yersinia en
6	42	64.6	130	1 YERA_YERPE	P31491 yersinia en
7	42	64.6	130	2 Q56910	Q56910 yersinia en
8	42	64.6	130	2 Q663P0	Q663P0 yersinia ps
9	42	64.6	669	2 Q98TU7	Q98TU7 agelatus ph
10	41	63.1	98	2 Q97FD1	Q97FD1 clostridium
11	41	63.1	322	2 Q23547	Q23547 caenorhabdi
12	41	63.1	456	2 Q98QW0	Q98QW0 mycoplasma
13	41	63.1	458	2 Q758Z3	Q758Z3 ashbya goss
14	41	63.1	1461	2 Q94H87	Q94H87 oryza sativ
15	40.5	62.3	111	2 Q9DB56	Q9DB56 mus musculus
16	40	61.5	261	2 Q9W5V6	Q9W5V6 drosophila
17	40	61.5	332	2 Q785D8	Q785D8 neurospora
18	40	61.5	374	2 Q75WT1	Q75WT1 gryllus bim
19	40	61.5	374	2 Q59420	Q59420 erwinia chr
20	40	61.5	469	2 Q8G7R9	Q8G7R9 bifidobacte
21	40	61.5	629	1 YD4B_SCHPO	Q10304 schizosacch
22	40	61.5	818	2 Q79CG2	Q79CG2 bacillus sp
23	40	61.5	824	2 Q45464	Q45464 bacillus sp
24	40	61.5	914	2 P93659	P93659 brassica na
25	40	61.5	988	2 P93065	P93065 brassica na
26	39.5	60.8	1190	2 Q68SS4	Q68SS4 pleurotus d
27	39	60.0	60	2 Q98IA7	Q98IA7 caenorhabdi
28	39	60.0	126	2 Q7NEX4	Q7NEX4 gloebacter
29	39	60.0	231	2 Q820Z3	Q820Z3 tropheryma
30	39	60.0	234	2 Q6NHQ5	Q6NHQ5 corynebacte
31	39	60.0	250	2 Q8N316	Q8N316 homo sapien

32	39	60.0	254	2 Q83N30	Q83N30 tropheryma
33	39	60.0	270	2 Q40987	Q40987 pieum sativ
34	39	60.0	276	2 Q8IW63	Q8IW63 homo sapien
35	39	60.0	285	2 Q8S860	Q8S860 oryza sativ
36	39	60.0	297	1 TRUA_DESVH	Q72418 desulfovibr
37	39	60.0	301	2 Q7XK8	Q7XK8 oryza sativ
38	39	60.0	305	1 LR25_HUMAN	Q8N386 homo sapien
39	39	60.0	305	2 Q6IQ00	Q6IQ00 homo sapien
40	39	60.0	314	2 Q66KC7	Q66KC7 xenopus tro
41	39	60.0	325	2 Q6GQ89	Q6GQ89 xenopus lae
42	39	60.0	347	2 Q96H00	Q96H00 homo sapien
43	39	60.0	357	2 Q98IA9	Q98IA9 caenorhabdi
44	39	60.0	394	2 Q9F2L1	Q9F2L1 streptomyce
45	39	60.0	402	2 P72675	P72675 synechocyst

ALIGNMENTS

RESULT 1
Q8N355 PRELIMINARY; PRT; 234 AA.
AC Q8N355;
DT 01-OCT-2002 (Tremblrel. 22, Created)
DE 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=23388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Boeak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC028090; AAH28090.1; -.
DR PIR; S09713; S09713.
DR PIR; S12441; S12441.
DR PIR; S30527; S30527.
DR HSP; P01842; ILIL.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003597; IG C1.
DR InterPro; IPR003006; IG MHC.
DR InterPro; IPR003596; IG LV.
DR Pfam; PF07654; Cl-set; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG LIKE; 2.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.

KW Hypothetical protein.
SQ SEQUENCE 234 AA; 24792 MW; CC948CAEBA4A9D63 CRC64;

Query Match 66.2%; Score 43; DB 2; Length 234;
Best Local Similarity 77.8%; Pred. No. 32;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 WDSGTHPV 11
|||||
DB 109 WDSGSDHEV 117

RESULT 2
Q96F86 PRELIMINARY; PRT; 508 AA.
AC Q96F86;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein FLJ21128 (Hypothetical protein PP844).
GN Name=FLJ21128; Synonyms=PP844;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye, and Placenta;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Diatchenko L., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska J., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RA Strausberg R.;
RN Submitted (JUL-2001) to the EMBL/GenBank/DBSJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Gu J.R., Wan D.F., Zhao X.T., Zhou X.M., Jiang H.Q., Zhang P.P.,
RA Qin W.X., Huang Y., Qiu X.K., Qian L.F., He L.P., Li H.N., Yu Y.,
RA Yu J., Han L.H.;
RN Submitted (OCT-1999) to the EMBL/GenBank/DBSJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Strausberg R.;
RN Submitted (JAN-2002) to the EMBL/GenBank/DBSJ databases.
RL EMBL; BC011534; AH11534.1; -;
DR EMBL; AF193058; AG22486.1; -;
DR EMBL; BC021271; RAH21271.1; -;
DR InterPro; IPR004443; YJef Nterm.
DR Pfam; PF03853; YJef N; 1.
KW Hypothetical protein.
SQ SEQUENCE 508 AA; 56077 MW; 95346F484FCF3EA CRC64;

Query Match 66.2%; Score 43; DB 2; Length 508;
Best Local Similarity 70.0%; Pred. No. 72;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 NSWDSSGTHP 10
|||||
DB 160 NSWSSSRHP 169

RESULT 3
Q9H797 PRELIMINARY; PRT; 508 AA.
AC Q9H797;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein FLJ21128.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,
RA Suzuki Y., Ohayashi M., Nishi T., Shibahara T., Tanaka T.,
RA Nakamura Y., Isoigaki T., Sugano S.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AK024781; BAB15001.1; -;
DR InterPro; IPR004443; YJef Nterm.
DR Pfam; PF03853; YJef N; 1.
SQ SEQUENCE 508 AA; 56105 MW; AAFABFD022E23DB CRC64;

Query Match 66.2%; Score 43; DB 2; Length 508;
Best Local Similarity 70.0%; Pred. No. 72;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 NSWDSSGTHP 10
|||||
DB 160 NSWSSSRHP 169

RESULT 4
Q9K2D3 PRELIMINARY; PRT; 508 AA.
AC Q9K2D3;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE AAS17853 protein (CDNA clone MGC:36552).
GN Name=Clk3; Synonyms=AAS17853;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J, and FVB/N; TISSUE=Mammary tumor;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Diatchenko L., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska J., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RA Strausberg R.;
RN Submitted (JUL-2001) to the EMBL/GenBank/DBSJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Gu J.R., Wan D.F., Zhao X.T., Zhou X.M., Jiang H.Q., Zhang P.P.,
RA Qin W.X., Huang Y., Qiu X.K., Qian L.F., He L.P., Li H.N., Yu Y.,
RA Yu J., Han L.H.;
RN Submitted (OCT-1999) to the EMBL/GenBank/DBSJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Strausberg R.;
RN Submitted (JAN-2002) to the EMBL/GenBank/DBSJ databases.
RL EMBL; BC011534; AH11534.1; -;
DR EMBL; AF193058; AG22486.1; -;
DR EMBL; BC021271; RAH21271.1; -;
DR InterPro; IPR004443; YJef Nterm.
DR Pfam; PF03853; YJef N; 1.
KW Hypothetical protein.
SQ SEQUENCE 508 AA; 56077 MW; 95346F484FCF3EA CRC64;

RA Krzywinski M.I., Skalska U., Smailus D.E., Schmerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CZECH II, and FVB/N; TISSUE=Mammary tumor;
 RA Strausberg R.;
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC031725; AAH31725.1; -;
 DR EMBL; BC033484; AAH33484.1; -;
 DR MGD; MGI:1098670; Clk3.
 DR InterPro; IPR004443; Yjef_Nterm.
 DR Pfam; PF03853; Yjef_N; 1_
 SQ SEQUENCE 508 AA; 55957 MW; DDIF449BEEF1341C8 CRC64;

Query Match 66.2%; Score 43; DB 2; Length 508;

Best Local Similarity 70.0%; Pred. No. 72;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 NSWSSGTHP 10

DB 160 NSWSSSRHP 169

RESULT 5

YERA_YEREN
 ID YERA_YEREN STANDARD; PRT; 130 AA.
 AC P31490;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE YopE regulator.
 GN Name-yerA;
 OS Yersinia enterocolitica.
 OG Plasmid pYV.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Yersinia.
 OX NCBI_TaxID=630;
 RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=8081 / Serotype O:8 / Biotype 1B;
 RX MEDLINE=90170873; PubMed=2307658;
 RA Forberg A., Wolf-Watz H.;
 RT "Genetic analysis of the yopE region of Yersinia spp.: identification
 RT of a novel conserved locus, yerA, regulating yopE expression.";
 RL J. Bacteriol. 172:1547-1555(1990).
 CC -!- FUNCTION: Positive regulator of yopE.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; M34278; AAA27673.1; -;
 DR PDB; IN5B; X-ray; A/B/C/D=1-130.
 DR InterPro; IPR010261; Cest.

DR InterPro; IPR005416; SycE_chap.
 DR Pfam; PF05932; Cest; 1

DR PRINTS; PR01596; SYCEHAPRONE.
 DR 3D-structure; Activator; Plasmid; Transcription regulation; Virulence.

SQ SEQUENCE 130 AA; 14676 MW; 491B67B4C293CASA CRC64;

Query Match 64.6%; Score 42; DB 1; Length 130;

Best Local Similarity 70.0%; Pred. No. 25;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 SWDSSGTHPV 11

|||||

DB 79 SWDEVGHPV 88

RESULT 6

YERA_YERPE
 ID YERA_YERPE STANDARD; PRT; 130 AA.
 AC P31491;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE YopE regulator.
 GN Name-yerA; OrderedLocusNames=YPCD1.05c, y5069, y0078, pCD82;
 OS Yersinia pestis.
 OG Plasmid pCD1, and Plasmid pYV019.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Yersinia.
 OX NCBI_TaxID=632;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=EV 76; PLASMID=pYV019;
 RX MEDLINE=90170873; PubMed=2307658;
 RA Forberg A., Wolf-Watz H.;
 RT "Genetic analysis of the yopE region of Yersinia spp.: identification
 RT of a novel conserved locus, yerA, regulating yopE expression.";
 RL J. Bacteriol. 172:1547-1555(1990).
 RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=KIMS / Biovar Mediaevalis; PLASMID=pCD1;
 RX MEDLINE=98427122; PubMed=9746557;
 RA Perry R.D., Straley S.C., Fetherston J.D., Rose D.J., Gregor J.,
 RA Blattner F.R.;
 RT "DNA sequencing and analysis of the low-Ca2+-response plasmid pCD1 of
 RT Yersinia pestis KIMS";
 RL Infect. Immun. 66:4611-4623(1998).
 RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=KIMS / Biovar Mediaevalis; PLASMID=pCD1;
 RX MEDLINE=98422474; PubMed=9748454;
 RA Hu P., Elliott J., McCready P., Skowronski E., Garnes J.,
 RA Kobayashi A., Brubaker R.R., Garcia E.;
 RT "Structural organization of virulence-associated plasmids of Yersinia
 RT pestis.";
 RL J. Bacteriol. 180:5192-5202(1998).
 RN [4]

RP SEQUENCE FROM N.A.

RC STRAIN=CO-92 / Biovar Orientalis; PLASMID=pCD1;
 RX MEDLINE=21470413; PubMed=11586360; DOI=10.1038/35097083;
 RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
 RA Prentice M.B., Sebahia M., James K.D., Churcher C.M., Mungall K.L.,
 RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.-M.,
 RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
 RA Feltwell T., Hamlin N., Holroyd S., Jagsis K., Karlyshev A.V.,
 RA Leather S., Moule S., Oyston P.C.F., Quail M.A., Rutherford K.M.,
 RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
 RT "Genome sequence of Yersinia pestis, the causative agent of plague.";
 RL Nature 413:523-527(2001).
 RN [5]

RP SEQUENCE FROM N.A.

RC STRAIN=91001 / Biovar Mediaevalis; PLASMID=pCD1;
 RA Song Y., Tong Z., Wang L., Han Y., Zhang J., Pei D., Wang J., Zhou D.,
 RA Han Y., Pang X., Zhai J., Chen H., Wang J., Li S., Guo Z.,
 RA Ye C., Du Z., Lin W., Wang J., Yu J., Yang H., Wang J., Huang P.,
 RA Yang R.;
 RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: Positive regulator of yopE.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

```
CC -----
DR EMBL; M34279; AAC27671.1; -
DR EMBL; AF074612; AAC69819.1; -
DR EMBL; AF053946; AAC62588.1; -
DR EMBL; AL117189; CAB54882.1; -
DR EMBL; AE017043; AAS58593.1; -
DR FIR; T43606; T43606.
DR PDB; 1JYA; X-ray; A/B=1-130.
DR PDB; 1K6Z; X-ray; A/B=1-130.
DR InterPro; IPR010261; CestT.
DR InterPro; IPR005416; Syce_chap.
DR Pfam; PF05932; Cest; 1.
DR PRINTS; PR01596; SYCECHAPRONE.
KW 3D-structure; Activator; Complete proteome; Plasmid;
KW Transcription regulation; Virulence.
FT HELIX 4 14
FT TURN 15 16
FT STRAND 26 31
FT TURN 32 33
FT STRAND 34 40
FT TURN 43 44
FT STRAND 45 50
FT TURN 56 57
FT HELIX 60 64
FT TURN 65 67
FT STRAND 73 74
FT TURN 77 81
FT STRAND 82 85
FT TURN 86 94
FT STRAND 95 97
FT TURN 100 101
FT HELIX 102 119
SQ SEQUENCE 130 AA; 14650 MW; 41A12BB29831CA5A CRC64;

Query Match 64.6%; Score 42; DB 1; Length 130;
Best Local Similarity 70.0%; Pred. No. 25;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 SWDSSGTHPV 11
DB 79 SWDEVGHPV 88
|||||

RESULT 7
ID Q56910 PRELIMINARY; PRT; 130 AA.
AC Q56910;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE YopE chaperone Syce.
GN Name=syCE;
OS Yersinia enterocolitica.
OG Plasmid pYve227.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=630;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=W22703;
RX MEDLINE=93268087; PubMed=8497188;
RA Wattiau P., Cornelis G.R.;
RT "Syce, a chaperone-like protein of Yersinia enterocolitica involved in
RT the secretion of YopE.";
RL Mol. Microbiol. 8:123-131(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=W22703;
RA Iriarte M., Lambermont I., Kerbouch C., Cornelis G.R.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF102990; RAD16849.1; -
DR HSP; P31490; INSE.
DR InterPro; IPR010261; CestT.

DR InterPro; IPR005416; Syce_chap.
DR Pfam; PF05932; Cest; 1.
DR PRINTS; PR01596; SYCECHAPRONE.
KW Plasmid.
SQ SEQUENCE 130 AA; 14649 MW; 490B89BA3293CA5A CRC64;

Query Match 64.6%; Score 42; DB 2; Length 130;
Best Local Similarity 70.0%; Pred. No. 25;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 SWDSSGTHPV 11
DB 79 SWDEVGHPV 88
|||||

RESULT 8
ID Q663P0 PRELIMINARY; PRT; 130 AA.
AC Q663P0;
DT 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE SYCE, verA: putative yopE chaperone.
GN ORFNames=PY0024;
OS Yersinia pseudotuberculosis IP 32953.
OG Plasmid pYV.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=273123;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IP 32953;
RX PubMed=15358858;
RA Chain P.S.G., Carniel E., Larimer F.W., Lamerdin J., Stoutland P.O.,
RA Regala W.M., Georgescu A.M., Vergez L.M., Land M.L., Motin L.V.,
RA Brubaker R.R., Fowler J., Hinnebusch B.J., Marceau M., Medigue C.,
RA Simonet M., Chenal-Francisque V., Souza B., Dacheux D., Elliott J.M.,
RA Derbise A., Hauser L.J., Garcia E.;
RT "Insights into the genome evolution of Yersinia pestis through whole
RT genome comparison with Yersinia pseudotuberculosis.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:13826-13831(2004).
DR EMBL; BX936399; CAF25367.1; -
DR InterPro; IPR010261; CestT.
DR InterPro; IPR005416; Syce_chap.
DR Pfam; PF05932; Cest; 1.
DR PRINTS; PR01596; SYCECHAPRONE.
KW Plasmid.
SQ SEQUENCE 130 AA; 14650 MW; 41A12BB29831CA5A CRC64;

Query Match 64.6%; Score 42; DB 2; Length 130;
Best Local Similarity 70.0%; Pred. No. 25;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 SWDSSGTHPV 11
DB 79 SWDEVGHPV 88
|||||

RESULT 9
ID Q98TU7 PRELIMINARY; PRT; 669 AA.
AC Q98TU7;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Serine kinase protein.
OS Agelaius phoeniceus (Red-winged blackbird).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Passeriformes; Icteridae; Agelaius.
OX NCBI_TaxID=39638;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=21365706; PubMed=11472064; DOI=10.1006/geno.2001.6596;
```

```

RA Gaeper J.S., Shiina T., Inoko H., Edwards S.V.;
RT "Songbird genomics: analysis of 45 kb upstream of a polymorphic Mhc
RL class II gene in red-winged blackbirds (Agelaius phoeniceus).";
DR EMBL; AF328738; AAK08500.1; -.
DR HSPF; Q13153; IF3M.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR Pfam; PF00069; Pkinase; I.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW Kinase.
SQ SEQUENCE 669 AA; 73710 MW; 5987154DCB1C441A CRC64;

Query Match 64.6%; Score 42; DB 2; Length 669;
Best Local Similarity 66.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 WDSGTHPV 11
|||
Db 170 WDEGHNHV 178

RESULT 10
Q97FD1 PRELIMINARY; PRT; 98 AA.
AC Q97FD1;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DE Predicted HD superfamily hydrolase.
GN OrderedLocusNames=CAC2809;
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1488;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VXM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RX DOI=10.1128/JB.183.16.4823-4838.2001;
RA Noelling J., Breton G., Omselchenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA Tatusov R.L., Sabathe F., Doucette-Stamm L.A., Soucaille P.,
RA Daly M.J., Bennett G.N., Koonin E.V., Smith D.R.;
RT "Genome sequence and comparative analysis of the solvent-producing
RT bacterium Clostridium acetobutylicum.";
RL J. Bacteriol. 183:4823-4838(2001).
DR EMBL; AE007778; AAK80753.1; -.
DR PIR; F97245; F97245
DR GO; GO:0016787; F:hydrolase activity; IEA.
KW Complete proteome; Hydrolase.
SQ SEQUENCE 98 AA; 11692 MW; 929C71F939AC930A CRC64;

Query Match 63.1%; Score 41; DB 2; Length 98;
Best Local Similarity 77.8%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 WDSGTHPV 10
|||
Db 40 WDSGTHPV 48

RESULT 11
Q23547

```

```

ID Q23547 PRELIMINARY; PRT; 322 AA.
AC Q23547;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein ZK596.3.
GN ORFNames=ZK596.3;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA McMurray A.A.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z69386; CAA93432.2; -.
DR PIR; T27931; T27931.
DR WormBase; WBGene00014008; ZK596.3.
DR WormPep; ZK596.3; CE35705.
KW Hypothetical protein.
SQ SEQUENCE 322 AA; 37172 MW; 224385DFB225AD80 CRC64;

Query Match 63.1%; Score 41; DB 2; Length 322;
Best Local Similarity 70.0%; Pred. No. 97;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 SWDSGTHPV 11
|||
Db 48 SWNSNGTHV 57

RESULT 12
Q98QW0 PRELIMINARY; PRT; 456 AA.
ID Q98QW0;
AC Q98QW0;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein MYPU_2500.
GN OrderedLocusNames=MYPU_2500;
OS Mycoplasma pulmonis.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2107;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=UAB CTIP;
RX MEDLINE=21267165; PubMed=11353084; DOI=10.1093/nar/29.10.2145;
RX Chamber I., Heilig R., Ferris S., Barbe V., Samson D., Galisson F.,
RA Moszer I., Dybvig K., Wroblewski H., Viari A., Rocha E.P.C.,
RA Blanchard A.;
RT "The complete genome sequence of the murine respiratory pathogen
RT Mycoplasma pulmonis.";
RL Nucleic Acids Res. 29:2145-2153(2001).
DR EMBL; AL445563; CAC13423.1; -.
DR PIR; B90543; B90543.
DR MypuList; MYPU_2500; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 456 AA; 53149 MW; 889D356B34946FB5 CRC64;

Query Match 63.1%; Score 41; DB 2; Length 456;
Best Local Similarity 70.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NSWDSSGTHP 10

```

```

Db      19 NSWDISKTYP 28
||||| | | |
RESULT 13
ID Q75823 PRELIMINARY; PRT; 458 AA.
AC Q75823;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE ADR384Wp.
GN ORFNames=ADR384W;
OS Asbya gossypii (yeast) (Eremothecium gossypii).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Eremothecium.
OX NCBI_TaxID=33169;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 10895;
RA Voegel S.E., Brachat S., Dietrich F.S., Lerch A., Gaffney T.,
RA Philippsen P.;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the cyclin family.
DR EMBL; AEO16817; AAS52304.1; -.
DR AGD; ADR384W; -.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR006670; Cyclin.
DR InterPro; IPR011028; Cyclin_like.
DR InterPro; IPR006671; Cyclin_N.
DR Pfam; PF00134; Cyclin_N; 1.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00385; CYCLIN; 1.
DR PROSITE; PS00292; CYCLINS; UNKNOWN_1.
KW Cyclin.
SQ SEQUENCE 458 AA; 50629 MW; C4A99D20DCE2BE65 CRC64;

Query Match 63.1%; Score 41; DB 2; Length 458;
Best Local Similarity 60.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 NSWSSSGTTP 10
| ||| : ||
Db 22 NWDSKASHP 31

RESULT 14
Q94H87 PRELIMINARY; PRT; 1461 AA.
AC Q94H87;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Putative disease resistance protein.
GN Name=OSUNBa0090P23.8;
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Buell C.R., Yuan Q., Ouyang S., Moffat K.S., Hill J.N., Gansberger K.,
RA Brenner M., Burgess S., Hance M., Shvartsbeyn M., Tsitrin T.,
RA Riggs F., Hsiao J., Zismann V., Blunt S., Pai G., Vanaken S.E.,
RA Utterback T.R., Feildlyum T.V., Quackenbush J., Salzberg S.L.,
RA White O., Fraser C.M.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Buell R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC084380; AAK52137.1; -.

DR Gramene; Q94H87; -.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR007090; LRR_plant.
DR InterPro; IPR002885; PPR.
DR InterPro; IPR008940; Prenyl_trans.
DR Pfam; PF00560; LRR_1; 18.
DR Pfam; PF01535; PPR; 5.
DR PRINTS; PR00019; LEURICHRPT.
DR TIGRFAMs; TIGR00756; PPR; 5.
SQ SEQUENCE 1461 AA; 158781 MW; 9A89A942B561D9F2 CRC64;

Query Match 63.1%; Score 41; DB 2; Length 1461;
Best Local Similarity 70.0%; Pred. No. 4.8e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 NSWSSSGTTP 10
| ||| | | |
Db 49 NSWPESGTSP 58

RESULT 15
Q9DB56 PRELIMINARY; PRT; 111 AA.
AC Q9DB56;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Mus musculus adult male cerebellum cDNA, RIKEN full-length enriched
DE library, clone:150010020 product:hypothetical protein, full insert
DE sequence.
GN Name=Kirrel3;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,

```

RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771 (2000).
RN [6]
RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Cerebellum;
RC Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK005197; BAB23877.1; -.
DR MGD; MGI:1914953; Kirrel3.
DR GO; GO:0005886; C:plasma membrane; IDA.
KW Hypothetical protein.
SQ SEQUENCE 111 AA; 12221 MW; 1BC7884BF056C43B CRC64;

Query Match 62.3%; Score 40.5; DB 2; Length 111;
Best Local Similarity 66.7%; Pred.No. 38;
Matches 8; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 1 NSWDSSG-THPV 11
||| :|||
Db 59 NSWRITGTPHV 70

Search completed: April 28, 2005, 18:24:20
Job time : 73.1562 secs

This Page Blank (uspto)